

Robert W. Boyd
Svetlana G. Lukishova
Victor N. Zadkov *Editors*

Quantum Photonics: Pioneering Advances and Emerging Applications

Springer Series in Optical Sciences

Volume 217

Founded by

H. K. V. Lotsch

Editor-in-chief

William T. Rhodes, Florida Atlantic University, Boca Raton, FL, USA

Series editors

Ali Adibi, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA

Toshimitsu Asakura, Hokkai-Gakuen University, Sapporo, Hokkaido, Japan

Theodor W. Hänsch, Max-Planck-Institut für Quantenoptik, Garching, Bayern, Germany

Ferenc Krausz, Garching, Bayern, Germany

Barry R. Masters, Cambridge, MA, USA

Katsumi Midorikawa, Laser Technology Laboratory, RIKEN Advanced Science Institute, Saitama, Japan

Bo A. J. Monemar, Department of Physics and Measurement Technology, Linköping University, Linköping, Sweden

Herbert Venghaus, Ostseebad Binz, Germany

Horst Weber, Berlin, Germany

Harald Weinfurter, München, Germany

Springer Series in Optical Sciences is led by Editor-in-Chief William T. Rhodes, Georgia Institute of Technology, USA, and provides an expanding selection of research monographs in all major areas of optics:

- lasers and quantum optics
- ultrafast phenomena
- optical spectroscopy techniques
- optoelectronics
- information optics
- applied laser technology
- industrial applications and
- other topics of contemporary interest.

With this broad coverage of topics the series is useful to research scientists and engineers who need up-to-date reference books.

More information about this series at <http://www.springer.com/series/624>

Robert W. Boyd · Svetlana G. Lukishova
Victor N. Zadkov
Editors

Quantum Photonics: Pioneering Advances and Emerging Applications



Springer

Editors

Robert W. Boyd
Department of Physics
University of Ottawa
Ottawa, ON, Canada

and

Institute of Optics and Department
of Physics
University of Rochester
Rochester, NY, USA

Svetlana G. Lukishova
The Institute of Optics
University of Rochester
Rochester, NY, USA

Victor N. Zadkov
Institute of Spectroscopy of the Russian
Academy of Sciences
Troitsk, Moscow, Russia

ISSN 0342-4111 ISSN 1556-1534 (electronic)
Springer Series in Optical Sciences ISBN 978-3-319-98400-1 ISBN 978-3-319-98402-5 (eBook)
<https://doi.org/10.1007/978-3-319-98402-5>

Library of Congress Control Number: 2018960218

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The word “quantum” became popular not only in physics, but in a much wider world: today’s general public is intrigued by extremely powerful quantum computers, quantum teleportation, and secure quantum communication in space. Usage of “quantum” in popular movies, TV shows, newspapers, and some consumer goods made it saturate the air for almost everybody, although often with the meanings different from those in physics.

The unique purpose of this book is to bring together the largely unreported history of experiments with single “light quanta” (photons) and map it with its modern manifestations as one of the key drivers to the vibrant scientific field of quantum photonics. We seek to convey to the readers the state-of-the-art advances of modern developments in quantum photonics and nonlinear optics from the leading groups throughout the world. The book also familiarizes the scientists, teachers, and students with the earliest experiments on single-photon and nonlinear optics from a time when both lasers and today’s light detectors did not yet exist, and the photographic plate or the human eye served as single-photon detectors. As a result, this book consists of two parts: (i) *Modern Quantum, Nano- and Nonlinear Photonics*, and (ii) *Historical Works: Single-Photon and Nonlinear Optical Experiments in the Pre-Laser Era* complimented with reprints and translation into English of some pioneering experimental papers. The book contains contributions of researchers from Australia, Canada, France, Germany, Hungary, Israel, Japan, Russia, Singapore, UK, and USA.

In this Preface, we clarify terminology and give relevant brief historical information. We also highlight the chapters of each of two parts of the book.

Terminology with Some Historical Highlights

Light Quanta (Photons)

In the third online edition of the Oxford English Dictionary (OED, December 2007), the first usage of the word “quantum” in English (borrowed from Latin) is dated back to 1567 (“... the body of Christe in the Sacramente is *quantum* ...”). Initially, this word had meaning of something that has quantity or total amount of quantity and was found chiefly in philosophy-related manuscripts or in law literature (with reference to amount of money).

According to the same edition of the OED, the usage of “quantum” in physics was recorded in 1870 (Nature **1**, p. 306) in a sense of quantity (“a certain quantum of the electric fluid”). Lord Kelvin also used “quantum” in a similar sense in both his 1902 and 1904 papers.

The modern meaning of the “quantum” concept in physics (a discrete quantity of electromagnetic energy proportional in magnitude to the frequency of the radiation, or any other physical discrete quantity (e.g., momentum or electric charge)) was born in the early twentieth century, owing to two classic German papers by Max Planck [1] and Albert Einstein [2]. We also would like to mention experiments of Millikan on the photoeffect confirming Einstein’s prediction [3] and Compton on inelastic X-ray scattering by electrons [4] after which the concepts of early quantum theory were accepted by the scientific community. See also the 1954 Nobel lecture by Glauber¹ on the history of light quanta [5].

Planck

Planck’s talk of December 14, 1900, and its publication [1] was later recognized as the turning point in the history of physics and currently is considered widely as the birth of quantum mechanics. Deriving oscillators’ entropy in the spectrum of blackbody radiation (distribution of a given energy among a set of oscillators producing the heat radiation emitted by black bodies), to satisfy experimental data, Planck introduced “energy elements” (Energieelemente, in German). He postulated that an energy E of the N resonators of frequency ν “to be composed of a very definite number of finite equal parts and we use thereto the constant of nature $h = 6.55 \times 10^{-27}$ erg s”. He continued further: “This constant multiplied by the common frequency ν of the resonators gives the energy element ϵ in erg, and dividing E by ϵ we get the number P of energy elements to be distributed over the N resonators” [1]; see also [6].

Einstein

Einstein was the first to recognize in 1905 that the discontinuity of Planck’s theory of blackbody radiation is not only a formal necessity for fitting the equation with the experimental data, but leads to new physics. Einstein started first the usage

¹ Roy Jay Glauber, one of the founding fathers of quantum optics and its “icon” through his quantum theory of optical coherence, widely cited in this book (Chaps. **1**, **3**, **4** and **19**) passed away while it went to print. Glauber will stay in our memory as Dr. Quantum Optics (coined by J. Eberly).

of the word “quanta” (plural) in the meaning of modern physics [2]: “Indeed, it seems to me that the observations regarding “blackbody radiation”, photoluminescence, production of cathode rays by ultraviolet light, and other groups of phenomena associated with the production or conversion of light can be understood better if one assumes that the energy of light is discontinuously distributed in space. According to the assumption to be contemplated here, when a light ray is spreading from a point, the energy is not distributed continuously over ever-increasing spaces, but consists of a finite number of energy quanta (Energiequanta, in German) that are localized in points in space, move without dividing, and can be absorbed or generated only as a whole.” In this paper [2], in the photoelectric effect section, he proposed a linear relationship between the maximum energy of electrons ejected from a surface and the frequency of the incident light. The slope of the line which did not depend on the substance was Planck’s constant. Verification of Einstein’s law for the energies of the photoelectrons came only in 1914–1916 by Millikan experiments, although even after them some doubts about the quantized nature of light persisted in scientific community, and only after Compton’s discovery in 1922 of X-ray quanta scattering by electrons obeying the same rules as the particle scattering, Einstein’s theory of light quanta was accepted.

Photon

The name “photon” is derived from Greek ($\varphi\omega\tauο$ = photo = light) and the “-on” at the end of the word was introduced in the twentieth century [7]. “Photon” as an alternative name for Einstein’s quantum of light became preferable by the mid-1930s [7]. Although it was used by several independent researchers earlier (see [7]), this word became synonymous with quantum of light only after the paper of Gilbert Newton Lewis of 1926 [8] in spite of its different meaning from Einstein’s light quantum: “... I therefore take the liberty of proposing for this hypothetical new atom, which is not light but plays an essential part in every process of radiation, the name photon” [8]. In our book, we have a chapter about Lewis’ contribution to nonlinear optics with a short biography of this outstanding scientist. According to [7], Arthur Holly Compton promoted the new term “photon.” In his 1927 Nobel lecture Compton wrote: “Here we do not think of the X-rays as waves but as light corpuscles, quanta, or, as we may call them, photons.” He also used it in popular literature.

Leonard Thompson Troland (famous American physicist, engineer and psychologist, president of the Optical Society of America (1922–1923)) first coined the word in 1916 (see [7] for more details and references therein), but he used it as a unit for the illumination of the retina. Five years later, it was independently introduced by the Irish physicist John Joly, professor of geology and mineralogy who was also interested in vision. What Joly called a “photon” was the minimum stimulus required to produce a signal in a fiber of the optic nerve. Then, in 1925, a French biochemist, biophysicist, and physiologist, René Wurmser (who was nominated twice for a Nobel Prize in Chemistry), in his paper about photochemical reactions relating to the role of chlorophyll in photosynthesis, wrote that “the activation of a molecule, in the sense of J. Perrin, demands the absorption of an

integral but variable number of photons.” Like Wurmser, but in a very different context, his compatriot, physicist Frithiof (Fred) Wolfers in his hypothesis also related to Perrin’s idea of molecular resonance (“*induction moléculaire*,” in French) suggested: “I shall use the name photons for the projectiles that supposedly transport radiant energy and possess the character of a periodic frequency ν (atoms of light). I then suggest that the photons may be repelled by the atoms of matter when they pass them closely … One may imagine that the repulsion is due to a kind of resonance between the photons and the resonators …” [7].

Photonics

In the current professional literature, “photonics” is used almost synonymously with the term “optics,” referring equally to both science and applications [9]. The phrase “optics and photonics” is used in [9] to capture light’s dual nature as:

- a propagating wave, like a radio wave, but with a much higher frequency;
- a collection of photons, with potential as a transformative field similar in impact to electronics.

According to the OED, the term “photonics” was mentioned already in 1952 in the Journal of the British Interplanetary Society: “From the fundamental domains of photonics, electronics, … and physical chemistry, our interest passes … to aerodynamics and the physics of solid bodies.” Independently, the term “photonics” (фотоника in Russian) was introduced in 1967 in the book in Russian [10] “Photonics of Molecules of Dyes and Related Compounds” of the Soviet academician Alexander N. Terenin. A European Commission CORDIS states that the term “photonics” was coined in 1967 by Pierre Aigrain [11], a prominent French physicist, a foreign member of the National Academy of Sciences (USA, since 1974): “Photonics is the science of the harnessing of light. Photonics encompasses the generation of light, the detection of light, the management of light through guidance, manipulation, and amplification, and most importantly, its utilization for the benefit of mankind.” In 1975, the book “Photonics” [12] (Proceedings of 1973–1974 years’ talks and a 1974 year conference) edited by M. Balkanski and P. Lallement was published in Paris in which this term was used by analogy with electronics, describing the application of the photon to the transmission of information, and included such topics as photon beam production, waveguiding, deflection, modulation, amplification, image processing, storage, and detection [13]. This book contains an explanatory note of Balkanski “Photonics: perspective in 1974” and Aigrain’s chapter in French “La photonique, technique de demain.” The term began to be seen in print in English around 1981 in press releases, annual reports of Bell Laboratories, and internal publications of Hughes Aircraft Corporation and in the more general press [9].

In the last decades, with development of integrated, nano-, biophotonics, free-space optical communication, and with the impact increasing of photonics on the national economies, the concept of “photonics” acquired a broader sense [9, 14]. It comprises also light–matter interaction, including nonlinear optical interaction

and extreme photonics. The twenty-first century will depend as much on photonics as the twentieth century depended on electronics (see the website of the International Year of Light (2015) [15]). In the book, we are using “photonics” in this broader sense as, for instance, in Ralf Menzel’s book, page 2 [16].

Part I. Modern Quantum, Nano- and Nonlinear Photonics

This part of the book consists of ten chapters and opens with Chap. 1 by Aspect and Grangier on the first single-photon sources and single-photon interference experiments. The authors emphasize the difference between “single-photon wave packets” and attenuated classical light pulses or light beams and describe the single-photon source developed by them in the mid-1980s (heralded single-photon wave packets, based on pairs of photons emitted in a radiative cascade). They also define the quantitative criterion of “anticorrelation” they use in distinguishing single-photon wave packets from attenuated pulses. In addition, the first single-photon interference experiment performed by the authors with their heralded source illustrates the notion of wave–particle duality. Brief overviews are also provided for both the first interference experiments in feeble light at 10^2 – 10^7 photons/s, and further developments in sources of single photons, heralded or on-demand, as well as in wave–particle duality experiments, in particular Wheeler’s delayed-choice experiment.

Mirhosseini, Lundeen, and Boyd (Chap. 2) provide an overview of recent progress in the tomography of structured light with an emphasis on the method known as direct measurement of the quantum wavefunction. Direct measurement provides a scalable and easy-to-implement approach for characterizing the transverse structure of single photons. This protocol is particularly attractive in light of the emerging role of high-dimensional optical states as a resource for encoding quantum information. This chapter presents a summary of various implementations of this technique that aim to characterize the spatial degree of freedom of the optical field.

Strelkov and Leuchs (Chap. 3) overview nonlinear interactions and non-classical light. This chapter includes the definition of non-classical light and basic examples and reviews some of the most prominent applications of non-classical light as well as the most common sources of non-classical light including physical systems of various sizes and complexity (ranging from single atoms to optical crystals and to semiconductor lasers). The authors also outline the trends in the field and the new cross-disciplinary approaches and techniques of generating non-classical light.

Lukishova and Bissell (Chap. 4) review room-temperature single-photon sources with photons exhibiting antibunching, including the authors’ results on single-photon sources with definite circular and linear polarizations. Single, “giant”, colloidal, semiconductor nanocrystal quantum dots and dot-in-rods, diamond color centers (both bulk and nanodiamonds), and trivalent rare-earth ions offer the best photostability (longest operating time) in room-temperature excitation. This review

highlights nanophotonic aspects of the problem, describing room-temperature single-photon sources based on these emitters and some new, stable, single emitters. Methods for emitter fluorescence enhancement (microcavities, including photonic bandgap, Bragg reflector and chiral liquid crystal microcavities, plasmonic nanoantennas, metamaterials), and the alignment of anisotropic single emitters with the help of liquid crystals are described and compared.

In Chap. 5, Victora, Kaneda, Bergmann, Wong, Graf, and Kwiat present an overview of time-multiplexing methods and time-bin qubits for quantum information processing with photons including applications of time-multiplexing techniques for more efficient single- and multiphoton sources, improved detectors, and high-bandwidth quantum memories, as well as enhanced applications such as quantum random walks and entanglement swapping. The results (by the authors) on experimental demonstration of a time-multiplexed, heralded, single-photon source are included.

Kravitsky and Volkov (Chap. 6) address the question of how fundamental photon fluctuations are perceived by a live visual system. The discussion is focused on photoreceptor cells within the eye, known as retinal rod cells. Rod cells provide vision under low-light conditions and are sensitive at a single-photon level. The authors review experiments on interaction of the rod cells with light sources of different photon statistics, including coherent, pseudo-thermal, and single-photon sources. Accurate control over photon statistics of light stimuli, combined with the technique for the readout of rod cells' response, enables precise and unambiguous characterization of intrinsic features of the visual system at single and discrete photon levels.

Chapter 7 by Fang, MacDonald, and Zheludev is devoted to controlling light with light via interference on planar photonic metamaterials. Planar photonic metamaterials—ultrathin media with nanoengineered optical properties—can realize the full potential of this concept to change optical data processing paradigms, spectroscopy, and nonlinear optics. Thin-film media can, if the film is much thinner than the light wavelength, lead to controllable energy exchange between incident and scattered waves and thereby to a plethora of new technological opportunities. This chapter describes how coherent interactions in metamaterials can facilitate nonlinear light-by-light control functions with THz bandwidth at arbitrarily low intensities.

Krasavin, Ginzburg, and Zayats (Chap. 8) review nonlinear plasmonics and plasmonic metamaterials. Plasmonics as a tool for tailoring and enhancing nonlinearity (coherent and Kerr-type nonlinearities, plasmonic metals as nonlinear materials), nonlinearities in plasmonic nanostructures, harmonic generation in plasmonic nanostructures, Kerr-type nonlinearity and ultrafast all-optical switching including nonlinear plasmonic crystals and optical bistability, nonlinear plasmonic metamaterials, epsilon-near-zero metamaterials, nonlinear surface plasmon polaritons are described in this chapter.

Makarov, in Chap. 9, concentrates on nonlinear optics with elliptically polarized singular beams and short pulses in media with spatial dispersion. The conditions of appearance and the behavior of polarization singularities in the cross section of a

light beam, arising due to nonlinear interaction of elliptically polarized laser beams with a medium with nonlocality of quadratic and cubic optical responses, are discussed. The formation dynamics and propagation features of polarization singularities, including pairwise creation and annihilation, for sum-frequency and second harmonic generation, beam self-action and interaction, and other nonlinear optical processes are presented. The author also discusses the effects accompanying the propagation of ultrashort (few oscillations) elliptically polarized light pulses in a nonlinear isotropic gyrotropic medium with frequency dispersion.

Zheltikov (Chap. 10) outlines ultrafast nonlinear optics in the mid-infrared. Recent breakthroughs in the generation of high-intensity, ultrashort laser pulses in the mid-infrared offer new approaches. This spectral range is unique in many ways, giving rise to new regimes of high-field nonlinear optics. Within this region, many molecular bands are located, drastically enhancing the coupling between the field and molecular motions. Electrons driven by intense field of ultrashort mid-infrared pulses acquire unusually high ponderomotive energies within a fraction of the field cycle. In addition, in this region the threshold power of self-focusing, proportional to square of the radiation wavelength, significantly increases. Consequently, much higher peak powers can be transmitted in a single laser filament in the mid-infrared range in comparison with the visible and near infrared, without losing beam continuity and spatial coherence.

Part II. Historical Works: Single-Photon and Nonlinear Optical Experiments in the Pre-Laser Era

This part of the book consists of nine chapters and contains both reprints or translations into English of experimental papers on the first quantum and nonlinear optical experiments during the pre-laser era as well as papers about the scientists who carried out these experiments (or their short biographies). The experiments of the researchers from different countries are included on the observation of the first interference fringes in a faint light, first light pressure measurements, first nonlinear optical experiments (saturation of absorption or luminescence) and their usage in practical devices, first experiments on sensitivity of a human eye to a faint light excluding physiological fluctuations smearing the results, first observation of statistical structure of interference field and independent fluctuations in two split coherent beams by a human eye in a faint light, and first photon-correlation measurements in split coherent beams using photomultipliers.

Chapter 11 by Lukishova includes Taylor's 1909 paper as well as his short biography. This paper is the first-reported experiment on interference fringes with very faint light. The light power in Taylor's experiment was 5×10^{-6} erg/s ($\sim 10^6$ photons/s) in a region of interference. Photographic plates were used for registration, and the maximum exposure time was about 3 months without changing the contrast of fringes. This experiment was suggested by Sir Joseph J. Thomson. Robert Millikan in his Nobel lecture [3] referred to the Thomson–Planck–Einstein's

concept of localized radiant energy (in its most general form introduced by Thomson in 1903).

The next two chaps. (12 and 13) contain papers of 1901 on the first experiments on measuring the light pressure, which were carried out independently and reported almost simultaneously by a Russian scientist, Lebedev (1900, August), and two Americans, Nichols and Hull (1901, August). These chapters begin with papers by Masalov about Lebedev's experiments and by Garmire about Nichols and Hull's experiments and the accuracy achieved at that time. The works of Lebedev and Nichols and Hull are frequently cited in the current scientific literature, e.g., on laser cooling and trapping as well as on cavity optomechanics.

Chapter 14 is devoted to Vavilov's contributions to photonics. It contains translations into English of parts of his two key papers: on the first nonlinear optical experiment (saturation of absorption in uranium glass at kW/cm^2 of spark light intensity, 1926) and on the sensitivity of the human eye to low-light level (1933). An original reprint of one of his papers in English of 1943 as a development of 1933 year work is also included. Using the human eye as a detector and flashes of faint light at eye's sensitivity threshold ($\sim 40\text{--}50$ "green" photons on the retina per flash in these measurements), Vavilov clearly observed independent intensity fluctuations of two split coherent beams, but sometimes two split beams were seen simultaneously (1933). He also studied a statistical structure of interference pattern. This chapter also contains the paper by Lukishova about Vavilov's contributions to photonics, specifically Vavilov-Čerenkov radiation (Vavilov was Čerenkov's thesis advisor, but had passed away when this effect deserved the 1958 Nobel Prize in physics).

Lewis, Lipkin, and Magel's 1941 paper on nonlinear optical effects (saturation in absorption and phosphorescence) is outlined in Chap. 15 by Lukishova. The authors investigated a fluorescein-doped boric-acid glass characterized by low saturation intensities. (In a recent publication [17] on the same material, the following numbers are reported: absorption saturation intensity of $\sim 15 \text{ mW/cm}^2$ and a nonlinear susceptibility $\chi^{(3)}$ as large as $\sim 1 \text{ esu}$, as compared to $\sim 10^{-12} \text{ esu}$ for the commonly used Kerr liquid CS_2). This chapter also contains Lewis' biography.

In Chap. 16 Stroud describes the nonlinear optical device of 1941 of the Institute of Optics, University of Rochester, based on saturation in luminescence. The Icaroscope by O'Brian was used by pilots during the Second World War. An original publication of O'Brian is reproduced.

In Chap. 17 by Lukishova, we included a reprint of the highly cited, 1941 paper by Hecht, Shlaer, and Pirenne on the sensitivity of the human eye at low-light level. Hecht et al. used a method similar to Vavilov's experiments of 1933–1942, and both groups independently arrived at similar results. In the same chapter, we also reprinted some excerpts from biography of Selig Hecht, member of the National Academy of Sciences (USA, since 1944).

The next two chaps. (18 and 19) are devoted to the first photon-correlation measurements using photomultipliers. Chapter 18 by Varró is devoted to the Hungarian scientist Jánossy who with his co-workers for the first time used photomultipliers and the photon-counting technique for photon-correlation measurements. Translation into English of paper of 1954 of Ádám, Jánossy and Varga is

added to this chapter. This part of their study is considered as the forerunner to the Hanbury Brown- and Twiss-type correlations with visible light, although Jánossy did not observe correlations in two split coherent beams. Chapter 18 begins with a biography of Jánossy. Chapter 19 by Lukishova and Tango is devoted to the Hanbury Brown–Twiss effect of 1956 on observation of positive correlations between two photomultiplier signals from a split-beam optical source. The history of its discovery is reprinted from the book of R. Hanbury Brown “Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics.” The short biographies of both researchers are also included. Tango who worked with Twiss wrote his biography for this book.

In conclusion, we believe that this book will be useful for academics, researchers, engineers, and students in many disciplines who hope to learn more about the history of quantum and nonlinear optics and upcoming trends in quantum photonics and nonlinear optics. Finally, we would like to thank all contributors who have found the time, energy, and enthusiasm to write these chapters. We highly appreciate the work of the translators of some papers into English. We thank American Institute of Physics, American Physical Society, Optical Society OSA, Cambridge Philosophical Society, the Rockefeller University Press, P. N. Lebedev Physical Institute, Russian Academy of Sciences (RAS), RAS Nauka Publisher (Moscow, Russia), S. I. Vavilov Institute for the History of Science and Technology of RAS (Moscow, Russia), Australian Physical Society, University of Rochester Library Archive, the Russian journal *Uspekhi Fizicheskikh Nauk* (Physics-*Uspekhi*), American Institute of Physics Emilio Segrè Visual Archives, Cavendish Laboratory (University of Cambridge), the Radio Society of Great Britain, Marion Hanbury Brown, and Norbert Kroo for help and permission in publishing images, journal reproductions, or text excerpts.

Rochester, NY, USA

Rochester, NY, USA/Ottawa, ON, Canada

Troitsk, Moscow, Russia

Svetlana G. Lukishova

Robert W. Boyd

Victor N. Zadkov

References

1. M. Planck, Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft* **2**, 237–245 (1900). English translation: M. Planck, On the theory of the energy distribution law of the normal spectrum, in *The Old Quantum Theory*, ed. by D. ter Haar (Pergamon Press, 1967), pp. 82–90
2. A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik* **17**, 132–148 (1905). English translation: A. Einstein, On a heuristic point of view about the creation and conversion of light, in *The Collected Papers of Albert Einstein*, Vol. 2, The Swiss Years: Writings, (1900–1909), ed. by J. Stachel, D.C. Cassidy, J. Renn, R. Schulmann, pp. 86–103. English translation by A. Beck and P. Havas, consultant, (Princeton University Press, 1989)

3. R.A. Millikan, The electron and the light-quant from the experimental point of view. *Nobel Lecture*, May 23 (1924) https://www.nobelprize.org/nobel_prizes/physics/laureates/1923/millikan-lecture.pdf
4. A.H. Compton, X-rays as a branch of optics. *Nobel Lecture*, December 12 (1927) https://www.nobelprize.org/nobel_prizes/physics/laureates/1927/compton-lecture.pdf
5. R.J. Glauber, One hundred years of light quanta, *Nobel Lecture*, December 8 (2005) https://www.nobelprize.org/nobel_prizes/physics/laureates/2005/glauber-lecture.pdf
6. H. Kragh, Max Planck: the reluctant revolutionary, *Physics World* **13**, 31–35 (2000)
7. H. Kragh, Photon: new light on an old name. <https://arxiv.org/abs/1401.0293>
8. G.N. Lewis, The conservation of photons. *Nature* **118**, 874–875 (1926)
9. *Optics and Photonics: Essential Technologies for our Nation* (The National Academies Press, Washington, DC 2013, <http://nap.edu/13491>)
10. A.N. Terenin, *Photonics of Molecules of Dyes and Related Compounds* (Moscow, Nauka 1967, in Russian)
11. http://cordis.europa.eu/fp7/ict/photonics/home_en.html
12. M. Balkanski, P. Lallemand (ed.), *Photonics* (Gauthier-Villars, Paris, 1975)
13. C.I. Colemann, Photonics, in *Physics Bulletin*, ed. by M. Balkanski, P. Lallemand (1976), March, p. 126.
14. National Photonics Initiative, <http://www.lightourfuture.org/>
15. The International Year of Light, <http://www.light2015.org/Home/WhyLightMatters/What-is-Photonics.html>
16. R. Menzel, *Photonics. Linear and Nonlinear Interactions of Laser Light and Matter*, Second edition, Springer (2007).
17. M.A. Kramer, W.R. Tompkin, R.W. Boyd, Nonlinear-optical interactions in fluorescein-doped boric acid glass. *Phys. Rev.* **34**, 2026–2031 (1986)

Contents

Part I Modern Quantum, Nano- and Nonlinear Photonics

1 The First Single Photon Sources and Single Photon Interference Experiments	3
Alain Aspect and Philippe Grangier	
1.1 A Paradoxical History	4
1.2 In Search of Single Photons	5
1.2.1 A Short History of Feeble Light Interference Experiments	5
1.2.2 An Attenuated Light Beam is Not Composed of Single-Photons	6
1.3 The Anticorrelation Test: A Criterion to Distinguish Single-Photon Wave Packets from Attenuated Pulses	8
1.4 Attenuated Classical Pulses Do Not Pass the Experimental Anticorrelation Test	12
1.5 Heralded Single-Photon Pulses Pass the Experimental Anticorrelation Test	13
1.5.1 Photon Pairs as a Resource for Heralded Single-Photons	13
1.5.2 The 1985 Anticorrelation Experiment	14
1.6 Single-Photon Interferences	17
1.6.1 Wave-Particle Duality in Textbooks	17
1.6.2 Interferences with a Single-Photon	17
1.7 Further Developments	20
1.7.1 Parametric Sources of Photon Pairs	20
1.7.2 Other Heralded and “On-Demand” Single-Photon Sources	20
1.7.3 “Delayed-Choice” Single-Photon Interference Experiments	21
References	21

2	Direct Measurement of the Photon's Spatial Wave Function	25
	Mohammad Mirhosseini, Jeff S. Lundeen and Robert W. Boyd	
2.1	Introduction	25
2.2	Strong and Weak Measurements	26
2.3	Direct Measurement of the Wave Function	30
2.4	Direct Measurement of the OAM State Vector	32
2.5	Compressive Direct Measurement	36
2.6	"Scan-Free" Direct Measurement	40
2.7	Comparison to Quantum State Tomography	43
2.8	Summary	45
	References	46
3	Nonlinear Interactions and Non-classical Light	51
	Dmitry V. Strekalov and Gerd Leuchs	
3.1	Introduction	51
3.1.1	Classical and Non-classical Light	51
3.1.2	Types of Non-classical Light	53
3.1.3	Applications of Non-classical Light	59
3.2	Sources of Non-classical Light	65
3.2.1	Atoms Real and Artificial	65
3.2.2	Parametric Down Conversion	70
3.2.3	Kerr Nonlinearity in Fibers and Resonators	79
3.2.4	Lasers and Other Feedback Systems	83
3.3	Final Remarks	85
	References	88
4	Nanophotonic Advances for Room-Temperature Single-Photon Sources	103
	Svetlana G. Lukishova and Luke J. Bissell	
4.1	Introduction	104
4.1.1	Photon Statistics of Different Light Sources	104
4.1.2	First Antibunching Experiments	106
4.1.3	Modern Trends on Development of Room-Temperature Single-Photon Sources	107
4.1.4	Single-Photon Generation and Characterization Units	109
4.1.5	Overview of Next Sections	110
4.2	Unbleachable Single Emitters for Room-Temperature Single-Photon Sources	111
4.2.1	"Giant" Colloidal Semiconductor Nanocrystal Quantum Dots and Dot-in-Rods	111
4.2.2	Color-Center Diamonds	113
4.2.3	Nanocrystals Doped with Trivalent Rare-Earth Ions	118
4.2.4	Other Types of Stable Single-Photon Emitters for Room-Temperature SPSs	121

4.3	Enhancement of Single-Photon Emission by Photonic and Plasmonic Structures	122
4.3.1	Collection Efficiency from a Single Emitter on a Cover Slip or Other Dielectric Surfaces	123
4.3.2	Single Emitters Within Microcavities Including Photonic Crystals	125
4.3.3	Single Emitters Within Plasmonic Nanoantennas	132
4.3.4	Single Emitters with Metamaterials	138
4.4	Single Emitters in Liquid Crystals: Single Photon Sources with Definite Circular and Linear Polarizations	141
4.4.1	Photonic Bandgap Chiral Microcavities in CLC	142
4.4.2	CLC Glassy Oligomeric Microcavities: NQD Fluorescence Enhancement and Antibunching	144
4.4.3	CLC Monomeric Microcavities: Circular Polarized Fluorescence and Photon Antibunching	146
4.4.4	Single-Emitter Alignment, Single-Photon Sources with Definite Linear Polarization, and Some Plasmonics Applications of Liquid Crystals	149
4.5	Fiber Integrated SPSs	154
4.6	Conclusion and Outlook	156
	References	159
5	Time-Multiplexed Methods for Optical Quantum Information Processing	179
	Michelle Victora, Fumihiro Kaneda, Fedor Bergmann, Jia Jun Wong, Austin Graf and Paul Kwiat	
5.1	Introduction	180
5.2	Time-Multiplexed Heralded Single-Photon Source	180
5.2.1	Introduction—Heralded Single-Photon Source	180
5.2.2	Working Principle	181
5.2.3	Experimental Realization	183
5.2.4	Enhanced Performance	186
5.3	Multi-photon Source	187
5.3.1	Multiple Photons in Multiple Modes	187
5.3.2	Fock States	189
5.4	Digital Quantum Memory	191
5.4.1	“Digital” Delay Line	191
5.4.2	Memory Cavity and Shutter Cavity	194
5.5	Applications	195
5.5.1	Time-Multiplexed Single-Photon Detectors	195
5.5.2	Time-Bin Qubits	196
5.5.3	Random Walks	198

5.5.4	Entanglement Swapping	200
References		202
6	The Role of Photon Statistics in Visual Perception	207
Leonid Krivitsky and Vadim Volkov		
6.1	Retinal Rod Cells	207
6.1.1	Morphology of Rod Cells	208
6.1.2	Light Detection by Rod Cells	209
6.1.3	Mechanisms of Reproducibility for Light-Induced Responses of Rod Cells	211
6.1.4	Comparison of Photoreceptors from Different Organisms and Man-Made Photodetectors	214
6.1.5	Basics of Communication Between Neurons and Rod Cells	215
6.1.6	Readout of Rod Cell Response	218
6.2	Overview of the Earlier Visual Experiments	219
6.3	Interaction of Rod Cells with Classical and Quantum Light Sources	222
6.3.1	Experiments with the Whole Visual System	223
6.3.2	Experiments with Isolated Rod Cells and Classical Light Sources	224
6.3.3	Interfacing Rod Cells with a True Single Photon Source	228
6.4	Conclusions and Future Perspectives	232
References		233
7	Controlling Light with Light via Interference on Photonic Metamaterials	239
Xu Fang, Kevin F. MacDonald and Nikolay I. Zheludev		
7.1	Introduction	239
7.1.1	Concepts and Applications of ‘Coherent Control’	239
7.1.2	Coherent Light-by-Light Control in Ultrathin Media	240
7.1.3	Effective Nonlinearity via Coherent Control	241
7.1.4	New Opportunities Presented by Nano-engineered Ultrathin Media	243
7.2	Coherent Interactions on Ultrathin Films	244
7.2.1	Singular Scattering Coefficients for Ultrathin Media	244
7.2.2	Effective Nonlinearity in Ultrathin Films	245
7.2.3	Scattering Coefficients for Real (Finite) Thin Film Media	246
7.2.4	Scattering Coefficients for Nano-engineered Metasurfaces	248

7.2.5	Scattering Coefficients for Lossless Media	250
7.2.6	Magnetic Field Scattering Coefficients	252
7.2.7	Cascading Metadevices	253
7.3	Coherent Metadevices	255
7.3.1	Small Signal Amplification	255
7.3.2	Logic Operations	256
7.4	Coherent Perfect Absorption in the Quantum Regime	261
7.5	Summary	261
	References	262
8	Nonlinear Nanoplasmonics	267
	Alexey V. Krasavin, Pavel Ginzburg and Anatoly V. Zayats	
8.1	Introduction	268
8.2	Brief Summary of Concepts of Plasmonics	269
8.2.1	Basic Notions of Plasmonics	269
8.2.2	Plasmonics as a Tool for Tailoring and Enhancing Nonlinearity	271
8.2.3	Plasmonic Metals as Nonlinear Materials	274
8.3	Nonlinearities in Plasmonic Nanostructures	275
8.3.1	Free-Electron Coherent Nonlinearities in Metals: Hydrodynamic Model	275
8.3.2	Kerr-Type Nonlinearity of Metals: Two-Temperature Model	278
8.4	Harmonic Generation in Plasmonic Nanostructures	280
8.4.1	Nonlinear Coupling of Plasmonic Resonances	281
8.4.2	SHG and THG from Metallic Nanostructures in Non-perturbative Hydrodynamic Description	284
8.4.3	Coherent Nonlinear Effects and Nonlocality	287
8.5	Kerr-Type Nonlinearity and Ultrafast Nonlinear Plasmonics	288
8.5.1	Controlling Light with Light	288
8.5.2	Controlling SPP Modes with Light	291
8.5.3	Nonlinear Plasmonic Crystals and Optical Bistability	293
8.5.4	Nonlinear Plasmonic Metamaterials	296
8.5.5	Epsilon-Near-Zero Metamaterials: Engineering Kerr Nonlinear Response	297
8.6	Nonlinear Surface Plasmon Polaritons	300
8.6.1	Nonlinear SPP Modes Due to Ponderomotive Nonlinearity	301
8.6.2	Plasmon-Solitons Due to Third-Order Nonlinearity	302
8.6.3	Cascaded Plasmon-Solitons Due to Second-Order Nonlinearity	303

8.7	Conclusion and Outlook	305
	References	307
9	Nonlinear Optics with Elliptically Polarized Singular Beams and Short Pulses in Media with Spatial Dispersion	317
	Vladimir A. Makarov	
9.1	Introduction	317
9.2	Elliptically Polarized Laser Beam and Polarization Singularities	319
9.2.1	Formal Characterization of Elliptically Polarized Laser Beam	319
9.2.2	Brief Information About Polarization Singularities in Laser Beams	321
9.2.3	Hodograph of the Electric Field Strength of Elliptically Polarized Ultrashort Laser Pulse	323
9.3	Nonlinear Electrodynamics of Homogeneous Media with Spatial Dispersion	325
9.3.1	Maxwell's Equations in Nonlinear Media with Spatial Dispersion	325
9.3.2	Material Equation in Nonlinear Media with Spatial Dispersion. Local and Nonlocal Susceptibilities	327
9.3.3	Boundary Conditions for Electromagnetic Field and Surface Nonlinear Optical Response. Surface Susceptibilities	329
9.3.4	Basis List of Nonlocal Nonlinear Optical Effects	331
9.3.5	Phenomenological Theory of Nonlinear Optical Response of the Bulk of Homogeneous Isotropic Media with Broken Mirror Symmetry	333
9.4	Polarization Singularities in the Electric Field at a Sum-Frequency Generated by Two Collinear Elliptically Polarized Beams in the Bulk of a Nonlinear Isotropic Medium with Spatial Dispersion of Nonlinear Optical Response	337
9.4.1	Formation of C-Type Polarization Singularities in the Beam Cross-Section at the Sum Frequency Generated by Uniformly Elliptically Polarized Gaussian Beams	337
9.4.2	Polarization Singularities in a Sum-Frequency Light Beam Generated by a Bichromatic Singular Beam	341
9.5	Singularities of the Sum-Frequency Light Field Polarization Arising upon Reflection on Incident Elliptically Polarized Gaussian Beams from the Surface of a Isotropic Medium with Spatial Dispersion of Nonlinear Optical Response	346
9.5.1	Nonuniform Transverse Distribution of the Light Intensity and Polarization upon Sum-Frequency Generation from the Surface	346

9.5.2	Polarization Singularities in Second Harmonic Beam Arising upon Reflection of Elliptically Polarized Gaussian Beam in the Case of Normal Incidence of Light	349
9.5.3	Polarization Singularities in Second Harmonic Beam Arising upon Reflection of Elliptically Polarized Gaussian Beam in the Case of Oblique Incidence	353
9.5.4	Singular Polarization Patterns of the Beam at Double Frequency Generated by Singarily Polarized Fundamental Beam	357
9.6	Formation and Interaction of Polarization Singularities During Light Self-Action and Interaction of Elliptically Polarized Beams in Isotropic Medium with Spatial Dispersion of Cubic Nonlinearity	364
9.6.1	Ring-Shaped Electric Field Structures Formation During Self-Focusing of Elliptically Polarized Gaussian Beam	364
9.6.2	Interaction of Two Monochromatic Beams with Polarization Singularities in Case of Their Collinear Propagation	369
9.7	Nonlinear Optical Activity in Propagation of Ultrashort Elliptically Polarized Laser Pulses in a Medium with Frequency Dispersion and Significant Spatial Dispersion of Cubic Nonlinearity	374
9.7.1	Physical Model of Third-Order Nonlinear Optical Susceptibility of Isotropic Medium with Frequency Dispersion and Significant Spatial Dispersion of Cubic Nonlinearity	374
9.7.2	Propagation of Ultrashort Elliptically Polarized Pulses in a Medium with Frequency Dispersion and Nonlocality of Nonlinear Optical Response	376
9.8	Conclusions	379
	References	380
10	Ultrafast Nonlinear Optics in the Mid-Infrared	385
	Alexei M. Zheltikov	
10.1	Introduction	386
10.2	Optical Physics in the Mid-Infrared: Why Longer Wavelengths Make a Difference	386
10.3	Ultrafast Optical Nonlinearities in the Mid-Infrared: The Role of Ionization Dynamics Within the Field Half-Cycle	390
10.4	Mid-Infrared Laser Filaments in the Atmosphere	395
10.5	Mid-Infrared-to-Mid-Ultraviolet Supercontinua	401

10.6	Multioctave Sub-Two-Cycle Supercontinua in the Mid-Infrared from Self-compressing, Self-focusing Soliton Transients in a Solid	401
10.7	Time-Domain Spectroscopy in the Mid-Infrared	404
10.8	Conclusion	412
	References	412
Part II Historical Works: Single-Photon and Nonlinear Optical Experiments in the Pre-Laser Era		
11	The First Paper on Experimental Observation of Interference Fringes with Feeble Light (Sir Geoffrey Ingram Taylor)	419
	Svetlana G. Lukishova	
11.1	Sir Geoffrey Ingram Taylor: Biography (Excerpts from Sir Brian A. Pippard's Paper in <i>Physics Today</i> 28, N9, 67 (1975))	421
11.2	Reprint of the First Paper on Experimental Observation of Interference Fringes with Feeble Light [from <i>Proceedings of the Cambridge Philosophical Society</i> 15, 114–115 (1909)]	422
12	First Experiments on Measuring Light Pressure I (Pyotr Nikolaevich Lebedev)	425
	Anatoly V. Masalov	
12.1	P. N. Lebedev—First Experiment on Measuring Light Pressure	425
12.2	“Experimental Examination of Light Pressure”. Original Paper of P. N. Lebedev on First Experimental Examination of Light Pressure	431
13	First Experiments on Measuring Light Pressure II (Ernest Fox Nichols and Gordon Ferrie Hull)	455
	Elsa Garmire	
13.1	Musings on the First Measurements of Light Pressure	455
13.1.1	Introduction	455
13.1.2	First Archival Reports of Light Pressure	456
13.1.3	Comparing Experiments	457
13.1.4	Toward Numerical Validation of Radiation Pressure	458
13.1.5	Mary Bell and S. E. Green Re-analysis of the Nichols and Hull Papers	460
13.1.6	Philosophers and Historians Have Their Say	461
13.1.7	Conclusions	462
13.2	Original Paper of E. F. Nichols and E. F. Hull “A Preliminary Communication on the Pressure of Heat and Light Radiation”	465

14 The First Nonlinear Optical Experiment of 1926, Measuring Sensitivity Threshold of the Human Eye to Feeble Light (1933) and Statistical Structure of Feeble-Light Interference by the Human Eye (Sergei Ivanovich Vavilov)	481
Svetlana G. Lukishova	
14.1 Sergei Vavilov's Nonlinear and Quantum Optical Experiments in Pre-laser Times, Vavilov-Čerenkov Radiation, and His Contributions to the Luminescence Science	482
14.1.1 Introduction	482
14.1.2 Vavilov's Nonlinear Optical Experiment (1926), Coining "Nonlinear Optics" Terminology and His Book "Microstructure of Light"	485
14.1.3 Method of Observation of Quantum Fluctuations of Light by a Human Eye (Threshold of Human-Eye Sensitivity) and Statistical Structure of Interference Pattern	490
14.1.4 Vavilov-Čerenkov Radiation	493
14.1.5 Vavilov's Contribution to the Luminescence Science and His School of Luminescence Which Created the Basis for First Soviet Lasers	495
14.1.6 Vavilov's Works on the History of the Science and Its Popularization	496
14.1.7 Vavilov's Organizing and Administrative Activities at the State Level	497
14.1.8 Some Biographical Information About S. I. Vavilov and His Family	500
14.2 The First Observation of Nonlinear Optical Effect (Saturation of Absorption): Translation of Excerpts from an Original Paper of S. I. Vavilov and V. L. Levshin "The Relation Between Fluorescence and Phosphorescence in Solid and Liquid Media"	508
14.2.1 The Possibility of Reducing the Absorption of Fluorescent and Phosphorescent Objects upon Their Illumination with Intense Light of a Spark (§ 4 of the Original Paper)	508
14.3 Visual Measurements of Statistical Fluctuations of Photons: Translation of an Original Paper by E. Brumberg and S. Vavilov	511
14.3.1 Introduction	511
14.3.2 The Theory of Experiment	513
14.3.3 Last Experiments	516
14.3.4 Results	519
14.3.5 Relative Fluctuations of Different Rays	527

14.3.6	Physiological and Photometric Consequences	529
14.3.7	Conclusions	531
14.4	S. I. Vavilov's Comments to the Publication of S. Hecht, S. Shlaer and M. H. Pirenne: Translation from His Book “The Microstructure of Light”	531
14.5	Reprint of an Original Paper by E.M. Brumberg, S. I. Vavilov and Z. M. Sverdlov “Visual Measurements of Quantum Fluctuations. I. the Threshold of Vision as Compared with the Results of Fluctuation Measurements”	534
15	Nonlinear Optical Experiment of 1941 (Gilbert Newton Lewis)	543
	Svetlana G. Lukishova	
15.1	Gilbert Newton Lewis: Biography	544
15.2	Nonlinear Optical Effects in Phosphorescence and Absorption (1941)	547
16	Nonlinear Optical Device (Icaroscope) at the Institute of Optics, University of Rochester During the Second World War	551
	Carlos R. Stroud	
16.1	A Brief History of Development of Icaroscope	551
16.2	Reprint of Original Publication About Icaroscope	553
17	Measuring Sensitivity Threshold of the Human Eye to Feeble Light (Selig Hecht)	555
	Svetlana G. Lukishova	
17.1	Selig Hecht: Biography	558
17.2	List of Selig Hecht Papers Related to Vision	560
17.3	Energy, Quanta and Vision (Original Paper)	564
18	Photon Correlation Measurements with Beamsplitter and Photomultipliers (Lajos Jánossy)	587
	Sándor Varró	
18.1	Lajos Jánossy: Biography	587
18.2	Coincidences of Photons Traveling in Coherent Beams of Light: Translation of an Original Paper by A. Ádám, L. Jánossy and P. Varga	590
18.2.1	Introduction	591
18.2.2	Accidental Coincidences and the Choice of the Coincidence Resolution Capability	591
18.2.3	The Method of the Measurement	594
18.2.4	The Optical Arrangement	594
18.2.5	The Counting Apparatus	596
18.2.6	Control Measurements	596
18.2.7	The Measurements	597

18.2.8	The Method of Data Processing	597
18.2.9	The Results of the Measurements	599
18.2.10	The Method of Frequent Changes	601
18.2.11	The Controlling Measurements	602
18.2.12	The Results of the Measurements	603
	References	604
19	First Observation of Photon Correlations (Bunching) with Beamsplitter and Photomultipliers (Robert Hanbury Brown and Richard Quintin Twiss)	605
	Svetlana G. Lukishova and William J. Tango	
19.1	Robert Hanbury Brown: Biography (1916–2002)	608
19.2	Richard Q. Twiss: Biography (1920–2005)	612
19.3	The Hanbury Brown–Twiss Effect	615
	Index	621

Editors and Contributors

About the Editors

Robert W. Boyd is Professor of Physics and Canada Excellence Research Chair Laureate in Quantum Nonlinear Optics at the University of Ottawa (<http://www.boydnlo.ca>) and Professor of Optics and Physics at the University of Rochester (<http://www.optics.rochester.edu/workgroups/boyd/>). He was born in Buffalo, New York. He received his B.S. in physics from MIT and the Ph.D. in physics from the University of California, Berkeley. His Ph.D. thesis was supervised by Charles Townes and involves the use of nonlinear optical techniques in infrared detection for astronomy. His research interests include studies of “slow” and “fast” light propagation, quantum imaging techniques, nonlinear optical interactions, studies of the nonlinear optical properties of materials, and the development of photonic devices including photonic biosensors. He has written two books, co-edited two anthologies, published over 400 research papers (\approx 39,000 citations, Google h-index 88), and been awarded ten patents. He is the recipient of the Arthur L. Schawlow Prize in Laser Science of the APS, the Charles Hard Townes Award of OSA, the Quantum Electronics Award of the IEEE Photonics Society, a Humboldt Research Award, and the Willis E. Lamb Award for Laser Science and Quantum Optics. He is a fellow of APS, OSA, IEEE, and SPIE. He is a past Chair of the Division of Laser Science of APS and has been a Member of the Board of Directors of OSA. He has served as a Member of the Board of Editors of Physical Review Letters and of the Board of Reviewing Editors of Science Magazine.

Prof. Svetlana G. Lukishova (<http://www.optics.rochester.edu/users/lukishov/>) is a Group Leader at the Institute of Optics, University of Rochester. Earlier she worked in Moscow, Russia, at the I. V. Kurchatov Nuclear Power Institute (Troitsk Branch) and the Institute of Radio-engineering and Electronics of the Russian Academy of Sciences and, after moving to the USA, at Liquid Crystal Institute (Kent, OH). She served six years as the Topical/Associate Editor of the journal “Optics Letters” in the field of nanophotonics, nonlinear optics, and liquid crystals. She is the author of more than 200 research publications (Physical Review Letters,

Optics Letters, Applied Physics Letters, etc.), and near 50 invited presentations in the field of quantum nanophotonics, nonlinear optics, and laser physics. Springer book “Self-focusing: Past and Present. Fundamentals and Prospects” (2009) edited by Lukishova (jointly with Boyd and Shen) is one of the top 25% most downloaded eBooks in the Springer eBook Collection in 2012 and 2015 [68 K chapter downloads during 2009–December 2018 with most downloaded chapter written by Lukishova (3130 downloads)]. She served as the Member of C.E.K. Mees Medal Committee of the Optical Society of America and as a Panelist for the National Science Foundation as well as an organizer/scientific board member of several conferences and symposia including International Quantum Electronics Conference, Optics of Liquid Crystals, Frontiers in Optics and others. She is the recipient of the W.C. Sykes Faculty Engineering Award of the University of Rochester and Diploma of 36th Vavilov’s Lectures on Luminescence of the Russian Academy of Sciences. She graduated with honors (red diploma) from the Department of General and Applied Physics of the Moscow Institute of Physics and Technology (MIPT) and defended her Ph.D. thesis (MIPT) in high-power laser physics working at the P.N. Lebedev Physical Institute of the Russian Academy of Sciences under the supervision of a Nobel Prize winner A. M. Prokhorov and P. P. Pashinin.

Prof. Victor N. Zadkov is a Director of the Institute of Spectroscopy of the Russian Academy of Sciences (since June 2015) also being a Professor of Physics with the Department of Physics, Lomonosov Moscow State University. Since the end of 2016, he is a Head of the Chair of Quantum Optics and Nanophotonics at the Faculty of Physics at the most prestigious non-classical university in Russia—Research University “Higher School of Economics.” He received his M.S., Ph.D., and Dr. Sci. in Physics from Lomonosov Moscow State University. His current research interests are in the fields of laser physics, interaction of laser radiation with matter, laser applications in life sciences, nanophotonics, quantum optics, and foundations of quantum information. He is an author of more than 150 refereed publications including a book and an editor of various collections and special issues of journals. He is a Fellow of the European Physical Society (EPS), in 2015 was elected as a Member of the EPS Council, and in 2018 as a Member of the EPS Executive Committee. He also served for many years as a jury member for the Prize on Research into the Science of Light, Fresnel Prizes, and QEOD Thesis Prizes. He is also a Member of the OSA and serve as a Member of the OSA International Council. He used to be a Member and then a Chair of the OSA Hopkins Leadership in Optics Award Committee for 2013–2016. In 2009, he was elected as a Member of the C-17 commission on Quantum Electronics of the IUPAP and in 2011—a Chair of this commission. Thanks to his efforts, the IUPAP gave a new life to the IUPAP Young Scientists Prize in Quantum Electronics and essentially enforced links with the EPS and QEOD/EPS. He used to be a member of the steering committee and member of the Russian committee for the International Year of Light 2015 and International Day of Light. He also stays in the roots of many collaboration projects in between Russia and Germany, France, Italy, Slovakia, USA, and China. They include running series of bilateral symposia on laser

physics, summer schools, organizing study tours, etc. Most successful are the ongoing German-Russian and French-Russian Laser Physics Symposia that resulted in lots of joint efforts. He serves on the editorial boards of the *European Physical Journal D*, *New Scientist* (RU), *Quantum Electronics*, *Moscow University Physics Bulletin* and as a program committee and international advisory board member for many international conferences and symposia. He also serves on many expert and research councils both national and worldwide. In 1984, he was awarded the Lenin Komsomol Prize in Physics (the highest award for young scientists in the Former Soviet Union). In 1997–1998, he used to work as a Research Fellow of the Alexander von Humboldt Foundation at the Institute of Applied Physics, University of Bonn, Germany. Since 2003, he is a President of the Moscow Humboldt Club. In 2008–2014, he served as a Humboldt Ambassador Scientist in Russia.

Contributors

Alain Aspect Laboratoire Charles Fabry, Institut d'Optique Graduate School, Université Paris-Saclay, Palaiseau, France

Fedor Bergmann Bergmann Messgeräte Entwicklung KG, Murnau, Germany

Luke J. Bissell Air Force Research Laboratory, Materials and Manufacturing Directorate, WPAFB, Dayton, OH, USA

Robert W. Boyd The Institute of Optics and Department of Physics, University of Rochester, Rochester, NY, USA; Department of Physics, University of Ottawa, Ottawa, ON, Canada

Xu Fang Optoelectronics Research Centre and Centre for Photonic Metamaterials, University of Southampton, Highfield, Southampton, UK

Elsa Garmire Thayer School of Engineering, Dartmouth College, Hanover, NH, USA; Santa Cruz, CA, USA

Pavel Ginzburg School of Electrical Engineering, Tel Aviv University, Tel Aviv, Israel

Austin Graf The Institute of Optics, University of Rochester, Rochester, NY, USA

Philippe Grangier Laboratoire Charles Fabry, Institut d'Optique Graduate School, Université Paris-Saclay, Palaiseau, France

Fumihiro Kaneda Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai, Japan

Alexey V. Krasavin Department of Physics and London Centre for Nanotechnology, King's College London, Strand, London, UK

Leonid Krivitsky Institute of Materials Research and Engineering, Agency for Science Technology and Research (A*STAR), Singapore, Singapore

Paul Kwiat Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Gerd Leuchs Max Planck Institute for the Science of Light, Erlangen, Germany

Svetlana G. Lukishova The Institute of Optics, University of Rochester, Rochester, NY, USA

Jeff S. Lundeen Department of Physics, University of Ottawa, Ottawa, ON, Canada

Kevin F. MacDonald Optoelectronics Research Centre and Centre for Photonic Metamaterials, University of Southampton, Highfield, Southampton, UK

Vladimir A. Makarov Faculty of Physics and International Laser Center, Lomonosov Moscow State University, Moscow, Russia

Anatoly V. Masalov Lebedev Physical Institute, Moscow, Russia

Mohammad Mirhosseini The Institute of Optics, University of Rochester, Rochester, NY, USA

Dmitry V. Strekalov Max Planck Institute for the Science of Light, Erlangen, Germany

Carlos R. Stroud The Institute of Optics, University of Rochester, Rochester, NY, USA

William J. Tango School of Physics, University of Sydney, Sydney, NSW, Australia

Sándor Varró Department of Theoretical Solid State Physics, Institute for Solid State Physics and Optics, Wigner Research Centre for Physics, Hungarian Academy of Sciences, Budapest, Hungary

Michelle Victoria Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Vadim Volkov Faculty of Life Sciences and Computing, London Metropolitan University, London, UK; Department of Plant Sciences, University of California, Davis, CA, USA

Jia Jun Wong DSO National Laboratories, Singapore, Singapore

Anatoly V. Zayats Department of Physics and London Centre for Nanotechnology, King's College London, Strand, London, UK

Alexei M. Zheltikov Department of Physics and Astronomy, Texas A&M University, College Station, TX, USA; Department of Physics, International Laser Center, M. V. Lomonosov Moscow State University, Moscow, Russia; Russian

Quantum Center, Moscow Region, Russia; Kurchatov Institute National Research Center, Moscow, Russia

Nikolay I. Zheludev Optoelectronics Research Centre and Centre for Photonic Metamaterials, University of Southampton, Highfield, Southampton, UK; Centre for Disruptive Photonic Technologies, School of Physical and Mathematical Sciences and The Photonics Institute, Nanyang Technological University, Singapore, Singapore

Part I

Modern Quantum,

Nano- and Nonlinear Photonics

Chapter 1

The First Single Photon Sources and Single Photon Interference Experiments



Alain Aspect and Philippe Grangier

Abstract This chapter shows how the concept of single-photon sources has emerged, in the mid 1980s. We emphasize the difference between “single-photon wave-packets” and attenuated classical light pulses or light beams. The quantum behavior of single photons—they cannot yield more than one photodetection—is contrasted with the behavior of attenuated classical light, which always yields some possibility of a joint detection on both sides of a beam splitter. We describe the single-photon source that we developed in the mid 1980s at Institut d’Optique, as well as the quantitative criterion (“anticorrelation”) that we introduced to show that it was indeed a single-photon source. We contrast these results with the ones that we obtained with a source of classical light pulses produced by a strongly attenuated light emitting diode, in which the average number of photons per pulse was much less than 1. We also describe the interference experiment we carried out with our single-photon source, illustrating the notion of wave-particle duality. We conclude with a brief overview of further developments in sources of single-photons, heralded or on-demand, as well as in wave-particle duality experiments, in particular Wheeler’s delayed choice experiments.

This chapter uses material already published,¹ where more detailed calculations can be found (see also [2]). A pedagogical presentation can be found in a MOOC centered on single photons.²

¹A. Aspect and P. Grangier, Chap. 10 in [1].

²<https://www.coursera.org/learn/quantum-optics-single-photon>.

A. Aspect (✉) · P. Grangier
Laboratoire Charles Fabry, Institut d’Optique Graduate School,
Université Paris-Saclay, Palaiseau, France
e-mail: alain.aspect@institutoptique.fr

1.1 A Paradoxical History

The rapidly developing field of quantum information [3] makes great use of two types of sources of quantum light: sources of single-photons on the one hand, sources of pairs of entangled photons on the other hand. One might think that single-photon sources were developed first, but it turns out that the history is just the opposite: in the optical domain, sources of pairs of entangled photons were developed first, and single-photon sources were invented later. This happened first with the 1972 source of entangled photons of Clauser and Freedman [4], about which a property related to the behavior of single-photons was demonstrated two years later [5]. In the same vein, it took five years for the efficient source of pairs of entangled photons of Aspect et al., designed for testing Bell's inequalities [6–8], to be explicitly used and characterized as the first source of single-photons by Grangier and Aspect [9]. Similarly, the first source of pairs of correlated photons produced by parametric down-conversion, demonstrated in 1970 by Burnham and Weinberg [10, 11], preceded by sixteen years the use of such a source to produce single-photons by Mandel et al. [12]. Actually, all these single-photon sources were what is called, in modern quantum optics jargon, “heralded single-photon sources”, i.e., sources that produce single-photon wave-packets whose leading-edge time—or peak time in the case of a bell-shaped pulse—is determined by the observation of the other photon of a pair [13]. This is why their development demanded the prior existence of a source of pairs of photons correlated in time. It took almost another two decades until the first source of single-photons “on-demand” appeared [14], i.e., a source of single-photon wave-packets whose leading-edge time can be chosen at will.

Although the question of single-photons had been repeatedly raised since the beginning of the 20th century, in the context of single-photon interference (Sect. 1.2.1), the question remained confused until the early-1980s, when we realized that none of the so-called “single-photon interference experiments” had been carried out with “one-photon states of light”. Indeed, all these experiments had been performed with feeble light beams issued from classical sources (such as discharge lamps), and it was clear, from the formalism of quantum optics recalled in Sect. 1.2.2, that as weak as they were, such light beams or light pulses could be described by quasi-classical states [15, 16]. Therefore, their properties could be understood by the semi-classical model of matter-light interaction, in which light is described as a classical electromagnetic wave, and the notion of a single-photon has no meaning. Inspired by the experiment of Clauser [5], and by the celebrated antibunching experiment of Kimble et al. [17], we discovered a simple quantitative criterion to test a characteristic property of a single-photon, *anticorrelation*: when sent onto a beam splitter, a single-photon (i.e., a one-photon Fock state of the quantized electromagnetic field) can be detected either on one side or on the other side of the beam-splitter, but never jointly on both sides. This is in contrast to the behavior of light that can be described by a classical wave, which, when split on the beam splitter, always yields a non-null rate of joint detection on both sides of the beam splitter. By demonstrating that this classical rate of joint detections has a minimum value, we obtained a quantitative

criterion that could be used to test the single-photon character of the light emitted by a source.

In Sect. 1.3 that criterion is derived, while Sect. 1.4 presents our experimental demonstration that strongly attenuated classical light pulses do not obey that criterion and thus do not pass the anticorrelation test.

Section 1.5 presents the idea of heralded single photon sources based on sources of photon pairs, describes our source, based on that principle, as well as the experimental results showing that the emitted pulses pass the anticorrelation test.

In Sect. 1.6, we give some details about the interference experiment we carried out with our single-photon source. Combined with the experiment of Sect. 1.5, which uses the same source, it yields a striking demonstration of the so-called wave-particle duality, one of the two great “mysteries of quantum mechanics” according to Feynman,³ and it can be used for an introductory course in quantum optics [2, 20, 21] (see also [22] and associated documents⁴).

Section 1.7 presents modern sources of single-photons, either heralded or on-demand, without details, since information on such sources can be found in [1] and in other chapters of the present book. We also mention further experiments on wave-particle duality, and in particular on Wheeler’s delayed choice experiment [23], which has been performed not only in its original form [24], but also in a more refined version [25, 26].

1.2 In Search of Single Photons

1.2.1 A Short History of Feeble Light Interference Experiments⁵

Almost as soon as Einstein introduced the notion of a quantum of light,⁶ i.e., a relativistic particle of energy $\hbar\omega$ and momentum $\hbar\omega/c$ [28], the question of the wave-like behaviour of the corresponding particle became a major concern among physicists, including Einstein himself [29]. The first attempt to investigate the question

³In his famous Lectures on Physics [18] Feynman cites Wave-Particle duality as “the only mystery” of quantum mechanics. However, two decades later [19] he emphasizes that there is a second quantum mystery, entanglement... and he immediately proposes to use it as a tool for quantum computing.

⁴A modern implementation of such an experiment illustrating wave-particle duality for a single-photon [22] has permitted our colleagues at ENS Cachan (ENS Paris-Saclay since 2017) to produce a video showing directly the construction of an interference pattern photon by photon, with a single-photon source passing the single-photon test. This video can be found for instance at <http://www.lcf.institutoptique.fr/Alain-Aspect-homepage>.

⁵The original papers describing some of these experiments, as well as other feeble light experiments, are reproduced in this volume.

⁶Einstein’s *LichtQuanten* was named photon only two decades later [27]. See also Chap. 5 of this volume.

experimentally [30] consisted of registering on a photographic plate the diffraction pattern of a needle illuminated with extremely attenuated light, so that the energy flux expressed in number of photons per second would correspond to an average distance between the photons much larger than the size of the apparatus. This pioneering experiment was followed by a long series of diffraction [31] and interference [32–37] experiments (see Table 1.1) with light emitted by strongly attenuated ordinary light sources, mostly discharge lamps, so that the average rate of photons entering the interferometric device, estimated as the power divided by the energy of a photon, ranged between 10^2 and 10^7 s^{-1} . Even at the largest of these rates, the average distance between photons was more than 10 m, much larger than the size of the interferometric device used in the corresponding experiment. It was thus concluded that there was only one-photon at a time in the interferometer, and the observation of fringes was then considered a demonstration that “a photon interferes with itself”. Actually, one experiment [34] failed to observe the interference pattern expected for a wave, but it was soon repeated by other scientists who found the expected interference pattern [36]. There is thus little doubt that diffraction or interference phenomena can be observed even in conditions of very weak light intensity.

In the 1970s, the general wisdom was then that single-photon wave-like behavior had been experimentally demonstrated. However, revisiting that question in the early 1980s, we realized that, according to the formalism of modern quantum optics as developed by Glauber [16, 38, 39], none of the experiments cited above could be considered a demonstration of single particle interference, because in none of these experiments the light used could be considered as composed of single-photon wave-packets. This lead us to perform the experiments of [9], presented in Sects. 1.5 and 1.6. In the rest of this section, we use the formalism of quantum optics to highlight the difference between single-photon wave-packets and all the types of light used in the experiments above.

1.2.2 An Attenuated Light Beam is Not Composed of Single-Photons

Quantization of one mode ℓ of the electromagnetic field leads to a Hamiltonian whose eigen-states $|n_\ell\rangle$ are eigenstates of the observable “number of photons in the mode ℓ ”:

$$\hat{N}_\ell = \hat{a}_\ell^\dagger \hat{a}_\ell , \quad (1.1)$$

the corresponding eigenvalue being precisely the number of photons n_ℓ :

$$\hat{N}_\ell |n_\ell\rangle = n_\ell |n_\ell\rangle . \quad (1.2)$$

Such a state is called “a number state”. A single-photon state is a number state associated with the value $n_\ell = 1$. Remarkably, the notion of number of photons

observable, and of single photon, can be generalized to a multimode situation. More precisely, the “total number of photons” observable is

$$\hat{N} = \sum_{\ell} \hat{N}_{\ell}, \quad (1.3)$$

while a state of the form

$$|1\rangle = \sum_{\ell} c_{\ell} |n_{\ell} = 1\rangle \quad (1.4)$$

is a one photon state (the notation $|n_{\ell} = 1\rangle$ implies that all modes other than ℓ are void, i.e., $|n_{\ell' \neq \ell} = 0\rangle$). One can indeed check that it is an eigenstate of \hat{N} (see 1.3) corresponding to the eigenvalue 1.

If now we consider an attenuated light beam, emitted for instance by a single mode c.w. laser, it is well known that the number of photons arriving on a detector during an interval of time T has a random character, described by a Poisson probability distribution

$$\mathcal{P}(N_{\ell}) = e^{-|\alpha_{\ell}|^2} \frac{(|\alpha_{\ell}|^2)^{N_{\ell}}}{N_{\ell}!} = e^{-\langle N \rangle} \frac{(\langle N \rangle)^{N_{\ell}}}{N_{\ell}!}. \quad (1.5)$$

The quantity

$$\langle N \rangle = |\alpha_{\ell}|^2 \quad (1.6)$$

is the average number of photons in the volume $S c T$, where S is the transverse surface of the beam and c the velocity of light. This distribution is associated with a quasi-classical state

$$|\alpha_{\ell}\rangle = e^{-|\alpha_{\ell}|^2/2} \sum_{n_{\ell}=0}^{\infty} \frac{\alpha_{\ell}^{n_{\ell}}}{\sqrt{n_{\ell}!}} |n_{\ell}\rangle, \quad (1.7)$$

which is an eigenstate of \hat{a}_{ℓ} with the eigenvalue α_{ℓ} :

$$\hat{a}_{\ell} |\alpha_{\ell}\rangle = \alpha_{\ell} |\alpha_{\ell}\rangle. \quad (1.8)$$

In a quasi-classical state, the number of photons is not well defined, and this remains true even if the average number of photons $\langle N \rangle$ is small compared to 1. In that case, the probability of finding one photon is

$$\mathcal{P}_{\langle N \rangle \ll 1}(1) \simeq \langle N \rangle \quad (1.9)$$

and the probability of finding two photons is

$$\mathcal{P}_{\langle N \rangle \ll 1}(2) \simeq \frac{\langle N \rangle^2}{2} \simeq \frac{[\mathcal{P}_{\langle N \rangle \ll 1}(1)]^2}{2}. \quad (1.10)$$

This property remains true if we consider an attenuated pulse, described as a multimode quasi-classical wave-packet

$$|\Psi_{qc}\rangle = |\alpha_{\ell=1}\rangle \otimes |\alpha_{\ell=2}\rangle \otimes \cdots \otimes |\alpha_{\ell}\rangle \otimes \cdots \quad (1.11)$$

with an average photon number

$$\langle N \rangle = \langle \Psi_{qc} | \hat{N} | \Psi_{qc} \rangle = \sum_{\ell} |\alpha_{\ell}|^2 \quad (1.12)$$

small compared to one.

We shall see in Sect. 1.3 that this entails dramatic differences with the case of a one-photon number state, and that the difference is yet larger when the classical light intensity fluctuates. In the case of chaotic light, either thermal light or light emitted by discharge lamps (see Sect. 10.2.7 in [1]), there is an extra factor of 2, associated with the Hanbury Brown and Twiss effect [40].

As explained in Chap. 10 of [1], classical continuous light beams emitted by a laser, a thermal source, or by a discharge lamp, are also described, in quantum optics, as a quasi-classical state, or a mixture (an incoherent superposition) of quasi-classical states. As a consequence, they cannot be considered as single photon wave packets, even if they are feeble enough that the wave packets do not overlap, in the average. It means that even in a time interval small enough that the average number of photons $P_{\langle N \rangle \ll 1}(1)$ is small compared to 1, there is a probability of having two photons at least equal to the value given by (1.10).

1.3 The Anticorrelation Test: A Criterion to Distinguish Single-Photon Wave Packets from Attenuated Pulses

The scheme of Fig. 1.1 allows one to unambiguously distinguish between single-photon pulses, and attenuated pulses from classical sources.⁷ It is based on the availability of a trigger signal, which generates an electronic gate activating photo-detectors during the time when the light pulse is expected to arrive. The pulse is launched onto a beamsplitter with two photo-detectors in the transmitted and reflected legs. If both detectors fire during the same gate, a coincidence is recorded. A counting system monitors the triggering events, the detection events at each photo-detector, and the coincidences.

A given experiment consists of running the source for a given duration, and counting the total number of counts in the transmitted (N_T) or reflected (N_R) channels, the total number of coincidences (N_c), and the total number of gates N_1 . We can then estimate the probabilities of single detections per gate,

⁷This scheme was used for the first time in the experiment [9] presented in Sect. 1.5. A related scheme was proposed independently in [41].

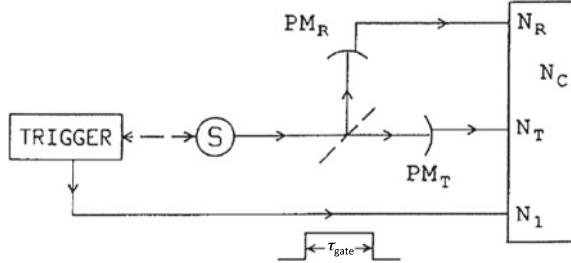


Fig. 1.1 Experiment to test the single-photon character of light pulses. The source S emits light pulses that fall on a beam splitter and can be detected in either channels (reflected, and transmitted) behind the beam splitter. The detectors are enabled during a gate τ_{gate} synchronized with the light pulses. The rates of single detection (N_R and N_T) and coincidence (N_C) are monitored. If the light pulse contains one-photon only, one detection is expected at most, and no coincidence is expected. In sources of heralded single-photons based on photon pairs, such as the ones emitted by the radiative cascade of Fig. 1.2, the trigger is activated by the detection of the other photon of the pair

$$P_R = \frac{N_R}{N_1} \quad \text{and} \quad P_T = \frac{N_T}{N_1}, \quad (1.13)$$

and the probability of a coincidence per gate,

$$P_c = \frac{N_c}{N_1}. \quad (1.14)$$

According to our intuition, we expect P_c to be zero in the ideal case of a one-photon wave-packet, and to be non-zero for a classical light pulse. We are going now to elaborate on this distinction in the context of the quantum theory of light.

For photo-detectors operated in the photon counting regime, fundamental quantities are the probabilities of single and double photo-detections. More precisely, for light in a state $|\Psi\rangle$, the probability of a photo-detection per unit of surface and time at the point \mathbf{r} and time t is

$$w^{(1)}(\mathbf{r}, t) = \eta \langle \Psi | \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) | \Psi \rangle, \quad (1.15)$$

where η is the sensitivity of the detector, and

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) = i \sum_{\ell} \mathcal{E}_{\ell}^{(1)} \epsilon_{\ell} \hat{a}_{\ell} \exp[i(\mathbf{k}_{\ell} \cdot \mathbf{r} - \omega_{\ell} t)], \quad (1.16)$$

while $\hat{\mathbf{E}}^{(-)}(\mathbf{r}, t)$ is the hermitian conjugate of $\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)$.

The probability of a double photo-detection per unit surface and time around \mathbf{r} and t and per unit surface and time around \mathbf{r}' and t' is

$$w^{(2)}(\mathbf{r}, t; \mathbf{r}', t') = \eta^2 \langle \Psi | \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \hat{\mathbf{E}}^{(-)}(\mathbf{r}', t') \hat{\mathbf{E}}^{(+)}(\mathbf{r}', t') \hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) | \Psi \rangle \quad (1.17)$$

The definitions of the various quantities involved in (1.15–1.17) can be found for instance in [2]. Here, the only important point is that the expressions (1.15) and (1.17) have the electric fields operators $\hat{\mathbf{E}}^{(-)}$ and $\hat{\mathbf{E}}^{(+)}$ appearing in the so-called “normal order”, i.e., with the destruction operators \hat{a}_ℓ on the right hand side of the creation operators \hat{a}_ℓ^\dagger .

In the case of a single-photon wave-packet, of the form (1.4), it has the dramatic consequence that

$$w^{(2)}(\mathbf{r}, t; \mathbf{r}', t') = 0 \quad (1.18)$$

since

$$\hat{a}_\ell |n_\ell = 1\rangle = |0\rangle, \quad (1.19)$$

where $|0\rangle$ is the vacuum, and

$$\hat{a}_\ell |0\rangle = 0. \quad (1.20)$$

To obtain the total probability of a joint detection on both sides of the beam splitter, one must integrate $w^{(2)}(\mathbf{r}, t; \mathbf{r}', t')$ over the surfaces of the detectors and over the time window associated with the gate, and the result is obviously

$$P_c = 0 \quad (1.21)$$

in agreement with our intuition. In contrast, the probabilities of single detections on either detector are different from zero, and they can be written as

$$P_R = \varepsilon_R R \quad \text{and} \quad P_T = \varepsilon_T T, \quad (1.22)$$

where ε_R and ε_T are the total quantum efficiencies of detection in each channel.

If now one has a classical pulse described by a quasi-classical state containing an average number of photons $\langle N \rangle$ small compared to 1, (1.22) is replaced by

$$P_R \simeq \varepsilon_R R \langle N \rangle, \quad P_T \simeq \varepsilon_T T \langle N \rangle, \quad (1.23)$$

and, because of (1.8), the probability of a coincidence is now

$$P_c \simeq \varepsilon_R \varepsilon_T R T \langle N \rangle^2, \quad (1.24)$$

i.e.,

$$P_c = P_R \cdot P_T. \quad (1.25)$$

This is in stark contradiction with (1.21–1.22). One has thus a striking difference between what is expected for a single-photon state and a quasi-classical pulse, even strongly attenuated. Note that the relation (1.25) can directly be obtained by considering the fraction of pairs of photons compared with the fraction of single photons in attenuated quasi-classical pulses, cf. (1.10).

One may wonder whether the contradiction is as strong if the average number of photons fluctuates over the various quasi-classical pulses. One must then replace $\langle N \rangle$ by $\overline{\langle N \rangle}$ in (1.23), where the overbar designates an ensemble average over the pulses. In the (1.24), $\langle N \rangle^2$ must be replaced by $\overline{\langle N \rangle^2}$, so that (1.25) must be replaced by

$$P_c \geq P_R \cdot P_T, \quad (1.26)$$

because of the Cauchy-Schwarz inequality.

$$\overline{x^2} \geq \bar{x}^2. \quad (1.27)$$

The contradiction between the case of a single-photon and the case of a quasi-classical pulse is yet stronger.

The criterion (1.26) can be expressed as a function of measured quantities. Using (1.13–1.14), we can write it as

$$\alpha = \frac{P_c}{P_R P_T} = \frac{N_c N_1}{N_R N_T} \geq 1. \quad (1.28)$$

This inequality holds for any classical light, even strongly attenuated, such as used in the experiments described in Table 1.1. Its violation is a clear indication that one

Table 1.1 Feeble light interference experiments. All these experiments have been realized with attenuated light from a classical source

Author	Date	Interferometer	Detector	Photon flux (s^{-1})	Interferences
Taylor [30]	1909	Diffraction	Photography	10^6	Yes
Dempster et al. [32]	1927	(i) Grating	Photography	10^5	Yes
		(ii) Fabry Perot	Photography	10^5	Yes
Janossy et al. [33]	1957	Michelson interf.	Photomultiplier	10^5	Yes
Donstov et al. [34]	1967	Fabry Perot	Image intensifier	10^3	No
Reynolds et al. [35]	1969	Fabry Perot	Image intensifier	10^2	Yes
Bozec et al. [36]	1969	Fabry Perot	Photography	10^2	Yes
Grishaev et al. [37]	1969	Jamin interf.	Image intensifier	10^3	Yes
Ciamberlini et al. [31]	1994	Diffraction	Image intensifier and CCD	10^5	Yes

has pulses with an indisputable single-photon character. We will consider this violation as a test to be passed by any light pulse claimed to have a single-photon character.

Remark

In the limit where the duration of the gate τ_{gate} is very small, the inequality (1.28) is equivalent to $g^{(2)}(0) \geq 1$, where $g^{(2)}(\tau)$ is the usual normalized second order correlation function [39]. So the condition $\alpha \geq 1$ can be seen as an “integrated” version of $g^{(2)}(0) \geq 1$, over a time window suited to the duration of the wave-packet. As for the semi-classical inequality $g^{(2)}(0) \geq 1$, or the semi-classical inequality $g^{(2)}(0) \geq g^{(2)}(\tau)$ used in [17], its violation has some relation with sub-Poissonian photon statistics, but no statistics is measured here, only intensity correlation functions: this is why we consider the wording “anticorrelation” well suited to characterize this violation.

1.4 Attenuated Classical Pulses Do Not Pass the Experimental Anticorrelation Test

To check our theoretical conclusion about attenuated classical pulses, and to test the photon counting system used in the experiments reported in [9], we applied the anticorrelation test to light from a pulsed LED (Light Emitting Diode). It produced light pulses with a rise time of 1.5 ns and a fall time about 6 ns. The gates, triggered by the electric pulses driving the photodiode, were 9 ns wide and had an almost complete overlap with the light pulses.

The source was attenuated to a level corresponding to one detection per 1,000 pulses emitted. With a detector quantum efficiency of about 10%, the average energy

Table 1.2 Anticorrelation experiment for light pulses from an attenuated photodiode (0.01 photon/pulse). The last column corresponds to the expected number of coincidences for $\alpha = 1$. All the measured coincidences are compatible with $\alpha = 1$: there is no evidence of anticorrelation. Note that the singles rates are similar to the ones of Table 1.3

Trigger rates	Singles rates		Duration	Measured coincidences	Expected coincidences for $\alpha = 1$
$N_1(s^{-1})$	$N_{2t}R(s^{-1})$	$N_{2f}T(s^{-1})$	$\theta(s)$	$N_c\theta$	$\frac{N_c N_T}{N_1} \theta$
4760	3.02	3.76	31,200	82	74.5
8880	5.58	7.28	31,200	153	143
12,130	7.90	10.2	25,200	157	167
20,400	14.1	20.0	25,200	341	349
35,750	26.4	33.1	12,800	329	313
50,800	44.3	48.6	18,800	840	798
67,600	69.6	72.5	12,800	925	955

per pulse can be estimated to be about 0.01 photon. In the context of Table 1.1, this source would have been considered a source of single-photons. The results presented on Table 1.2 show that it is definitely not the case. The quantity α is consistently found equal to 1; i.e., no anticorrelation is observed. In fact, the coincidence rate is exactly in agreement with the marginal case of inequality (1.28). This means that the pulses emitted by the LED had very little fluctuations.

This experiment thus supports the claim that light emitted by an attenuated classical source does not exhibit one-photon behavior on a beamsplitter, even in the case of very attenuated light pulses with an average energy by pulse much less than the energy of a photon.

1.5 Heralded Single-Photon Pulses Pass the Experimental Anticorrelation Test

1.5.1 *Photon Pairs as a Resource for Heralded Single-Photons*

When an atom is raised to an excited level, it re-emits a one-photon wave-packet, as can be guessed merely from energy conservation. However, in usual sources, such as discharge lamps, many excited atoms are seen simultaneously by the detectors, and their times of excitation are random. The theoretical description of such light is thus a mixture of one-photon wave-packets with random leading-edge times (Sect. 10.2.7 in [1]). If one also takes into account the fluctuations of the number of excited atoms, the emitted light can be considered a statistical ensemble of quasi-classical states, and there is no hope in this situation to observe any non-classical effect.

To observe the non-classical properties of Sect. 1.3 in fluorescent light, the most natural solution is to isolate single-atom emission, either in space, or in time.⁸ Isolating single-photon wave-packets in time is achieved in sources of “heralded” single-photons, based on the emission of *separated pairs of photons*: the first photon then “heralds” the emission of the second one, allowing one to isolate single-atom emission in time.

The production of pairs of photons occurs in many different contexts in physics, including particle physics (e.g., electron-positron annihilation, producing two γ photons), nuclear physics and atomic physics (through cascade de-excitation between several levels), and non-linear optics (pair production in spontaneous parametric fluorescence). In the latter case, the temporal correlation between the two photons of one pair, a fully quantum property, was observed first in 1970 by Burnham and Weinberg [10] and studied more accurately by Hong et al. [44], while some specifically

⁸It does not mean that non classical effects cannot be observed in many atoms fluorescence, as shown for instance in [42] or [43].

quantum properties of the photon pairs emitted by an atomic cascade were demonstrated in 1974 by Clauser [5].

However, it took some time to realize that a very simple way to understand these specifically quantum properties is to consider the quantum state of the second photon only, once the first one has been detected: according to the “projection postulate” of quantum mechanics, this second photon is then in a state very close to a one-photon state, more precisely a one-photon wave-packet with a well defined leading-edge time (or peak time). One can say that the second photon is “heralded” by the detection of the first one [13]. This wording has been adopted by the quantum optics community, and is now used as a generic name for such sources.

1.5.2 The 1985 Anticorrelation Experiment [9]

We developed an experiment corresponding to the scheme of Fig. 1.1, i.e., a setup allowing us to measure the single and coincidences rates on the two sides of a beam-splitter, during the opening of gates triggered by events synchronous with the light pulses. This system was used to study light pulses from a source designed to emit heralded one-photon wave-packets, based on pairs of photons emitted in a radiative cascade, as presented in this section, but also used to study pulses emitted by a classical source and strongly attenuated (Sect. 1.4).

Our source, initially designed to produce pairs of entangled photons to test Bell’s inequalities [6–8], was composed of atoms excited to the upper level of a two-photon radiative cascade (Fig. 1.2). Each excited atom decays by emission of two photons at different frequencies ν_1 and ν_2 . The time intervals between the detections of ν_1 and

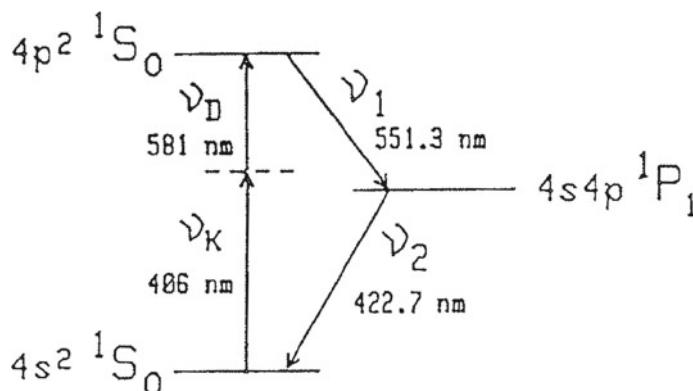


Fig. 1.2 Radiative cascade in ^{40}Ca calcium atoms, used to produce heralded single-photon pulses. The atom is excited to the upper level of the cascade by a resonant two-photon excitation with a Krypton ion laser and a tunable dye laser. It then re-emits photons ν_1 and ν_2 . Detection of photon ν_1 activates the trigger of Fig. 1.1

v_2 are distributed according to an exponential law, corresponding to the decay time of the intermediate state (lifetime $\tau_s = 4.7$ ns, which is also the time constant Γ^{-1} of the wave-packet describing the heralded single-photon v_2). By choosing the rate of excitation much smaller than $(\tau_s)^{-1}$, we have cascades well separated in time, in the average. We use the detection of v_1 as a trigger for a gate of duration $\tau_{\text{gate}} = 2\tau_s$, corresponding to the scheme of Fig. 1.1. During a gate opening, the probability of detecting a photon v_2 coming from the atom that emitted v_1 is much larger than the probability of detecting a photon v_2 coming from any other atom in the source. We are then in a situation close to an ideal single-photon pulse, and we expect to observe the corresponding anticorrelation behavior as predicted in Sect. 1.3.

The expected values of the counting rates can be obtained from a straight-forward quantum mechanical calculation. Denoting N the rate of excitation of the cascades, and η_1 , η_T and η_R the detection efficiencies of photons v_1 and v_2 (including the collection solid angles, optics transmissions, and detector efficiencies) we obtain:

$$N_1 = \eta_1 N \quad (1.29)$$

$$N_T = N_1 \eta_T [f(\tau_{\text{gate}}) + N\tau_{\text{gate}}] \quad (1.30)$$

$$N_R = N_1 \eta_R [f(\tau_{\text{gate}}) + N\tau_{\text{gate}}] \quad (1.31)$$

$$N_c = N_1 \eta_T \eta_R [2f(\tau_{\text{gate}})N\tau_{\text{gate}} + (N\tau_{\text{gate}})^2], \quad (1.32)$$

where $N\tau_{\text{gate}}$ is the probability to have a photon from another atom than the heralding atom, during the gate. The quantity $f(\tau_{\text{gate}})$, very close to 1 in this experiment, is the product of the factor $[1 - \exp(-\tau_{\text{gate}}/\tau_s)]$ (overlap between the gate and the exponential decay) and a factor somewhat greater than 1 that is related to the angular correlation between v_1 and v_2 [6, 9].

The quantum mechanical prediction for α is

$$\alpha_{\text{QM}} = \frac{2f(\tau_{\text{gate}})N\tau_{\text{gate}} + (N\tau_{\text{gate}})^2}{[f(\tau_{\text{gate}}) + N\tau_{\text{gate}}]^2}, \quad (1.33)$$

which is smaller than 1, as expected. The anticorrelation effect will be strong (α small compared to 1) if $N\tau_{\text{gate}}$ is much smaller than 1, i.e., the average number of cascades excited during a gate is much smaller than 1. This condition is easily fulfilled if the cascades are well separated in time, in the average.

Counting electronics, including the gating system, was a critical part of this experiment. The gate τ_{gate} was realized by logical decisions based on the measurement of the time intervals between the detections at the various detectors. This allowed the adjustment of the gates with an accuracy of 0.1 ns. The system also yielded various time-delay spectra, useful for consistency checks.

Table 1.3 shows the measured counting rates for different values of the excitation rate of the cascade. The corresponding values of α have been plotted in Fig. 1.3 as a function of $N\tau_{\text{gate}}$. As expected, the violation of inequality (1.28) increases as $N\tau_{\text{gate}}$ decreases, but the signal decreases simultaneously, and it becomes necessary

Table 1.3 Anticorrelation experiment with single-photon pulses from the radiative cascade. The last column corresponds to the expected number of coincidences for $\alpha = 1$. The measured coincidences show a clear anticorrelation effect. These data can be compared to Table 1.2.

Trigger rates	Singles rates		Duration	Measured coincidences	Expected coincidences for $\alpha = 1$
$N_1(s^{-1})$	$N_R(s^{-1})$	$N_T(s^{-1})$	$\theta(s)$	$N_c\theta$	$\frac{N_R N_T}{N_1} \theta$
4720	2.45	3.23	1200	6	25.5
8870	4.55	5.75	17,200	9	50.8
12,100	6.21	8.44	14,800	23	64.1
20,400	12.6	17.0	19,200	86	204
36,500	31.0	40.6	13,200	273	456
50,300	47.6	61.9	8400	314	492
67,100	71.5	95.8	3600	291	367

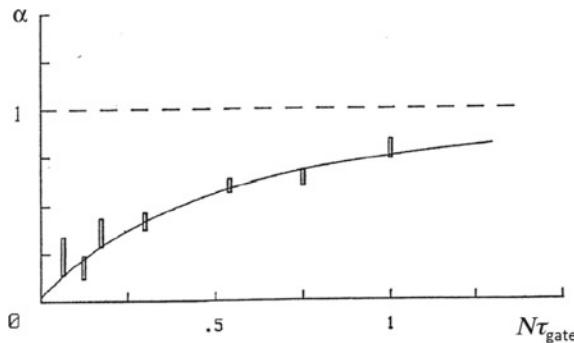


Fig. 1.3 Correlation parameter α as a function of the excitation rate N . The value of α smaller than 1 is the signature of an anticorrelation, corresponding to one-photon behaviour (for classical pulses, even strongly attenuated, one always have $\alpha \geq 1$). The solid line is the prediction of quantum optics taking into account the possibility that more than one atom is excited during one gate of duration τ_{gate} .

to accumulate the data for long enough periods of time to achieve a reasonable statistical accuracy. A maximum violation of more than 13 standard deviations has been obtained for a counting time of five hours (second line of Table 1.3). The value of α then is 0.18(6), corresponding to a total number of coincidences of 9, instead of the minimum value of 50 expected for a quasi-classical pulse.

1.6 Single-Photon Interferences

1.6.1 Wave-Particle Duality in Textbooks

Many introductory courses in Quantum Mechanics—whether or not they choose an historical perspective—begin with the description of an “experiment” exhibiting the wave-particle duality of for material particles. This is usually presented with an interference pattern, obtained for instance with Young’s slits. Such a phenomenon is then interpreted by invoking a wave that passes through both holes: it is well known that the resulting intensity then depends on the “path difference” Δ , and exhibits a modulation depending on the interference order $p = \Delta/\lambda$, where λ is the wavelength. On the other hand, the “particle” character is usually considered obvious for matter particles such as electrons, neutrons or atoms.

In the case of light, as discussed in previous sections, the particle character is questionable, unless one uses single photon wave-packets. This is why it is necessary to first test the single photon character of the light emitted by the source, before using the same light to test the ability to produce interference fringes. As we will see in this section, the combination of these two experiments constitutes a spectacular illustration of wave-particle duality.

1.6.2 Interferences with a Single-Photon

The quantum theory of light predicts that interferences are to be expected even with one-photon pulses (see for instance [2] for a detailed calculation). We have thus built a Mach-Zehnder interferometer, keeping the same source and the same beamsplitter as in Fig. 1.1, but removing the detectors on both sides of the beam splitter, and recombining the two beams on a second beam splitter (Fig. 1.4) [9]. The detection rates in the two outputs (1) and (2) are expected to be modulated as a function of the

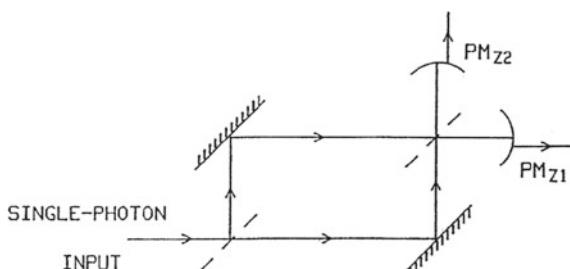


Fig. 1.4 Single-photon interference experiment. The source and the beamsplitter are similar to Fig. 1.1, but are now configured as a Mach-Zehnder interferometer. The detectors are gated as in Fig. 1.1, synchronously with the light pulses

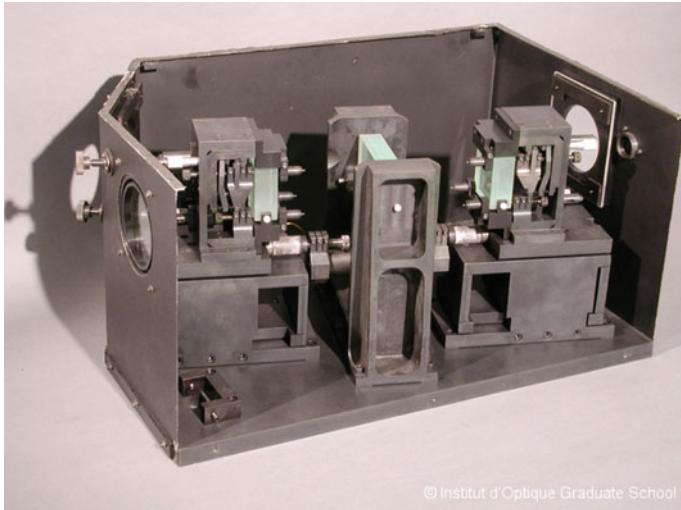


Fig. 1.5 Mach-Zehnder interferometer used in the single-photon interference experiment of 1985 [9]. The left and right thick plates hold multielectric coatings realizing the two mirrors, while the thick plate in the middle, twice as large, holds the two beam-splitting coatings. The sophisticated displacement mechanics is actuated by the two PZT (white cylinders). The input is at the front left round aperture, and the outputs are at the right round aperture and at the back left aperture, hidden but whose shadow can be seen

path difference in both arms of the interferometer. To guarantee that we are still working with one-photon pulses, the detectors PM_1 and PM_2 are gated synchronously with the pulses, as they were in the experiment of Sect. 1.5.1.

The interferometer (Fig. 1.5) was carefully designed and built to give high-visibility fringes with the beam of large étendue (product of transverse area by solid angle) produced by our source (about $0.5 \text{ mm}^2 \text{ rad}^2$). The reflecting mirrors and the beam splitters were $\lambda/50$ flat over a 40 mm diameter aperture. A mechanical system driven by piezoelectric transducers permitted displacement of the mirrors while keeping their orientation exactly constant: this allowed us to perfectly control the path difference of the interferometer. Preliminary checks with classical light showed a strong modulation of the counting rates of PM_{Z1} and PM_{Z2} when the path difference was modified. For a source shaped as the one-photon pulses source, the measured visibility was $V = 98.7(5)\%$, a value very close to the ideal value $V = 1$, showing the high quality of the interferometer.

Figure 1.6 presents the results obtained by running this interferometer with the one-photon source. The numbers of counts during a given time interval were measured at different values of the path difference. In the first plots, the counting time at each position was 0.01 s, while it was 10 s for the last recordings. This run was performed with the sources at a regime corresponding to an anticorrelation parameter $\alpha = 0.2$, definitely in the one-photon regime. These recordings clearly show the interference fringes building up “one-photon at a time”. When enough data have been

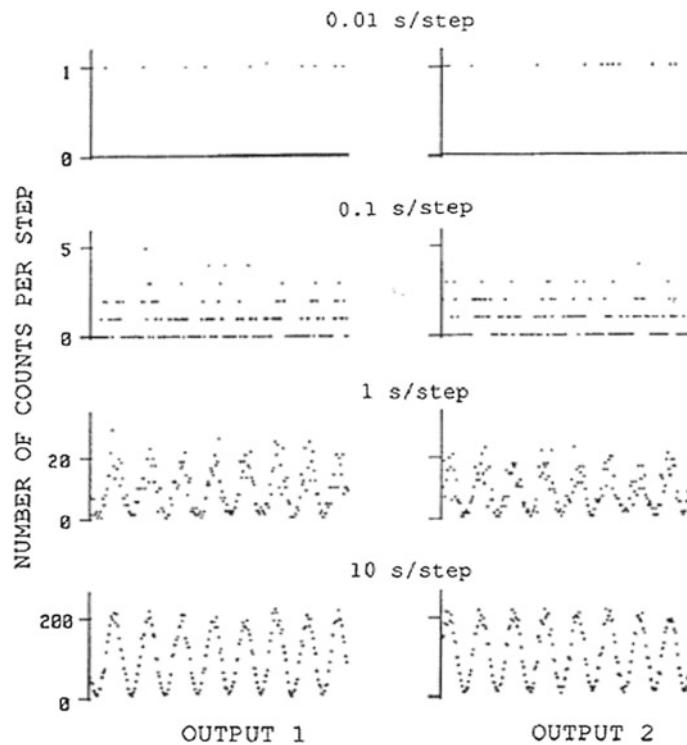


Fig. 1.6 Number of detected counts in output (1) and (2) as a function of the path difference. The four sets of plots correspond to different counting times at each path difference. This experiment has been realized in the single-photon regime ($\alpha = 0.2$). Note that the interferograms of outputs (1) and (2) are complementary

accumulated, the signal to noise ratio is high enough to allow a measurement of the visibility of the fringes. We repeated such measurements for various regimes of the source, corresponding to the different values of α shown on Fig. 1.3, and observed no deviation from the expected value $V = 98.7$, within the experimental noise, even in a regime where the source emits almost pure one-photon pulses. As predicted by the quantum theory of light, single-photon pulses do interfere. To our knowledge, this experiment (realized in 1985) was the first of this kind to be performed with a “fully quantum” light source, for which the anticorrelation effect was also directly observed.

1.7 Further Developments

1.7.1 Parametric Sources of Photon Pairs

During the same period as the experiments described above—that is between the early 1970s and the mid 1980s—another approach to generating photon pairs was developed using Parametric Down Conversion (PDC) in $\chi^{(2)}$ non-linear crystals, rather than atomic cascades [10, 11]. In 1986, Hong and Mandel performed an experiment strongly related to the anticorrelation effect described above, though presented in a different way [12]. The non-classical features of these photon pairs are similar to the ones of pairs produced in atomic radiative cascades as the one of Sect. 1.5.2, with some noticeable differences:

- Due to phase-matching conditions, parametric photons are strongly correlated both in their emission times, with a sub-nanosecond time separation of the order of the inverse of the phase-matching bandwidth, and in their emission directions, due to the conservation of the photon momenta when “splitting” a pump photon into two parametric photons. As a consequence of the latter, the heralded photons can be collected with an efficiency orders of magnitude better than in an atomic cascade, and this has been intensively used in experiments.
- From a practical point of view, a PDC experiment is significantly simpler and more reliable than an atomic cascade experiment. Nowadays, a photon anticorrelation experiment is a small and simple table-top experiment, which can be operated by students in laboratory classes.⁹

For these reasons, PDC is now widely used to produce heralded single-photons, and it is even possible to produce number states in well-defined spatio-temporal modes, and to reconstruct their Wigner functions using quantum homodyne tomography. This has been demonstrated both for one-photon [45] and two-photon Fock states [46]. It can be noted also that photon pairs can be emitted from $\chi^{(3)}$ non-linearities in optical fibers, rather than $\chi^{(2)}$ non-linearities in crystals, as described for instance in Chap. 12 of [1].

1.7.2 Other Heralded and “On-Demand” Single-Photon Sources

Many other types of single-photons sources have been proposed and implemented, using quantum dots, single molecules or single atoms, possibly in the cavity QED regime, Nitrogen-Vacancy centers in diamond, collectively enhanced quantum ensembles. Some of these sources are “on-demand” single-photon sources, meaning

⁹See for instance <http://www.institutoptique.fr/Formation/Ingenieur-Grande-Ecole/Travaux-Pratiques/Physique-quantique-atomique-nanophysique> at Institut d’Optique Graduate School.

that the single-photon is not only “heralded”, i.e., produced at a time that we can know, but emitted in a “push-button” way, at a time chosen by the experimentalist. This can be obtained rather easily from pulsed excitation of a single quantum emitter. In addition it is desirable that the photon is emitted with a very high efficiency (that is, each “click” gives one and only one-photon), in a perfectly defined spatio-temporal mode. The fully ideal, 100% efficient, on-demand single-photon source is not available yet, but impressive progress has been achieved¹⁰ since our 1986 year experiment.

1.7.3 “Delayed-Choice” Single-Photon Interference Experiments

To conclude, let us mention some more recent developments in single-photon interferences. Following a celebrated proposal by Wheeler [23], a very convincing “delayed-choice” interference experiment has been performed by Jacques et al., using a Nitrogen-Vacancy (NV) center in diamond as the single-photon source [24]. In this experiment, the “choice” of leaving the interferometer open—and thus observing the “which path” information¹¹—or closing the interferometer—and thus observing the interference fringes—was made while the photon was already inside a 50-m long interferometer. In more recent experiments, it was shown that this choice can be made remotely, by using a second photon entangled with the photon inside the interferometer [25, 26]. These experiments are a few examples only, which demonstrate the impressive control that can be obtained in manipulating single-photons, offering more and more possibilities for applications in quantum information and quantum communications.

References

1. A. Migdall, S.V. Polyakov, J. Fan, J.C. Bienfang, *Single-Photon Generation and Detection: Physics and Applications*, vol. 45 (Academic Press, 2013)
2. G. Grynberg, A. Aspect, C. Fabre, *Introduction to Quantum Optics: From the Semi-Classical Approach to Quantized Light* (Cambridge University Press, 2010)
3. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010)
4. S.J. Freedman, J.F. Clauser, Experimental test of local hidden-variable theories. Phys. Rev. Lett. **28**(14), 938–900 (1972)
5. J.F. Clauser, Experimental distinction between quantum and classical field-theoretic predictions for photoelectric effect. Phys. Rev. D **9**(4), 853–860 (1974)

¹⁰See for instance [43], Chap. 13 of [1], or Chaps. 3 and 4 of the current book.

¹¹It should be emphasized that in such experiments, the incompatible quantities are the which-path distinguishability and the visibility of the fringes (see [47–49] and week 6 in <https://www.coursera.org/learn/quantum-optics-single-photon>).

6. A. Aspect, P. Grangier, G. Roger, Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**(7), 460–463 (1981)
7. A. Aspect, P. Grangier, G. Roger, Experimental realization of Einstein-Podolsky-Rosen-Bohm gedanken experiment—a new violation of Bell inequalities. *Phys. Rev. Lett.* **49**(2), 91–94 (1982)
8. A. Aspect, J. Dalibard, G. Roger, Experimental test of bell inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49**(25), 1804–1807 (1982)
9. P. Grangier, G. Roger, A. Aspect, Experimental-evidence for a photon anticorrelation effect on a beam splitter—a new light on single-photon interferences. *Europhys. Lett.* **1**(4), 173–179 (1986)
10. D.C. Burnham, D.L. Weinberg, Observation of simultaneity in parametric production of optical photon pairs. *Phys. Rev. Lett.* **25**(2), 84–000 (1970)
11. S. Friberg, C.K. Hong, L. Mandel, Measurement of time delays in the parametric production of photon pairs. *Phys. Rev. Lett.* **54**(18), 2011–2013 (1985)
12. C.K. Hong, L. Mandel, Experimental realization of a localized one-photon state. *Phys. Rev. Lett.* **56**(1), 58–60 (1986)
13. D.T. Pegg, R. Loudon, P.L. Knight, Correlations in light emitted by 3-level atoms. *Phys. Rev. A* **33**(6), 4085–4091 (1986)
14. B. Lounis, W. Moerner, Single photons on demand from a single molecule at room temperature. *Nature* **407**(6803), 491–493 (2000)
15. E.C.G. Sudarshan, Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Phys. Rev. Lett.* **10**, 277–279 (1963)
16. R.J. Glauber, Coherent and incoherent states of the radiation field. *Phys. Rev.* **131**(6), 2766 (1963)
17. H.J. Kimble, M. Dagenais, L. Mandel, Photon anti-bunching in resonance fluorescence. *Phys. Rev. Lett.* **39**(11), 691–695 (1977)
18. R.P. Feynman, *Lectures on Physics* (Addison-Wesley, 1963)
19. R.P. Feynman, Simulating physics with computers. *Int. J. Theor. Phys.* **21**(6–7), 467–488 (1982)
20. R. Loudon, *The Quantum Theory of Light* (Oxford University Press, 2000)
21. C. Gerry, P. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2005)
22. V. Jacques, E. Wu, T. Tourny, F. Treussart, A. Aspect, P. Grangier, J.F. Roch, Single-photon wavefront-splitting interference—an illustration of the light quantum in action. *Eur. Phys. J. D* **35**(3), 561–565 (2005)
23. J.A. Wheeler, *Quantum Theory and Measurement* (Princeton University Press, 1984)
24. V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, J.F. Roch, Experimental realization of wheeler's delayed-choice gedanken experiment. *Science* **315**(5814), 966–968 (2007)
25. F. Kaiser, T. Coudreau, P. Milman, D.B. Ostrowsky, S. Tanzilli, Entanglement-enabled delayed-choice experiment. *Science* **338**(6107), 637–640 (2012)
26. A. Peruzzo, P. Shadbolt, N. Brunner, S. Popescu, J.L. O'Brien, A quantum delayed-choice experiment. *Science* **338**(6107), 634–637 (2012)
27. G.N. Lewis, The conservation of photons. *Nature* **118**, 874–875 (1926)
28. A. Einstein, Generation and conversion of light with regard to a heuristic point of view. *Annalen Der Physik* **17**(6), 132–148 (1905)
29. A. Einstein, On the evolution of our vision on the nature and constitution of radiation. *Physikalische Zeitschrift* **10**, 817–826 (1909)
30. G.I. Taylor, Interference fringes with feeble light. *Proc. Camb. Philos. Soc.* **15**, 114–115 (1910)
31. C. Ciamberlini, G. Longobardi, Real-time analysis of diffraction patterns, at extremely low-light levels. *Opt. Lasers Eng.* **21**(5), 317–325 (1994)
32. A.J. Dempster, H.F. Batho, Light quanta and interference. *Phys. Rev.* **30**(5), 644–648 (1927)
33. L. Jánossy, Z. Náray, The interference phenomena of light at very low intensities. *Acta Physica Academiae Scientiarum Hungaricae* **7**(4), 403–425 (1957)
34. Y.P. Dontsov, A.I. Baz, Interference experiments with statistically independent photons. *Sov. Phys. JETP-USSR* **25**(1), 1–000 (1967)

35. G.T. Reynolds, K. Spartalian, D.B. Scarl, Interference effects produced by single photons. *Nuovo Cimento Della Societa Italiana Di Fisica B-General Physics Relativity Astronomy and Mathematical Physics and Methods* **61**(2), 355 (1969)
36. P. Bozec, M. Cagnet, G. Roger, Experiments on interference in a weak light. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences Serie B* **269**(18), 883–000 (1969)
37. I.A. Grishaev, N.N. Naugolny, L.V. Reprints, A.S. Tarasenk, A.M. Shendero, Interference experiment and photon statistics for synchrotron radiation from electrons in a storage ring. *Sov. Phys. JETP-USSR* **32**(1), 16 (1971)
38. R. Glauber, *Quantum Optics and Electronics. Les Houches Summer School 1964*, ed. by C. DeWitt (Gordon and Breach, New York, 1965)
39. R.J. Glauber, The quantum theory of optical coherence. *Phys. Rev.* **130**(6), 2529 (1963)
40. R.H. Brown, R. Twiss, Correlation between photons in two coherent beams of light. *Nature* **177**(4497), 27–29 (1956)
41. B. Saleh, M. Teich, Sub-Poisson light generation by selective deletion from cascaded atomic emissions. *Opt. Commun.* **52**(6), 429–432 (1985)
42. P. Grangier, G. Roger, A. Aspect, A. Heidmann, S. Reynaud, Observation of photon antibunching in phase-matched multiatom resonance fluorescence. *Phys. Rev. Lett.* **57**(6), 687–690 (1986)
43. E. Bimbard, R. Boddeda, N. Vitrant, A. Grankin, V. Parigi, J. Stanojevic, A. Ourjoumtsev, P. Grangier, Homodyne tomography of a single photon retrieved on demand from a cavity-enhanced cold atom memory. *Phys. Rev. Lett.* **112**(3), 033601 (2014)
44. C.K. Hong, Z.Y. Ou, L. Mandel, Measurement of subpicosecond time intervals between 2 photons by interference. *Phys. Rev. Lett.* **59**(18), 2044–2046 (1987)
45. A.I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, S. Schiller, Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87**(5), 050402 (2001)
46. A. Ourjoumtsev, R. Tualle-Brouri, P. Grangier, Quantum homodyne tomography of a two-photon Fock state. *Phys. Rev. Lett.* **96**(21) (2006)
47. W.K. Wootters, W.H. Zurek, Complementarity in the double-slit experiment: quantum non-separability and a quantitative statement of Bohr's principle. *Phys. Rev. D* **19**(2), 473 (1979)
48. B.-G. Englert, Fringe visibility and which-way information: an inequality. *Phys. Rev. Lett.* **77**(11), 2154 (1996)
49. V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, J.F. Roch, Delayed-choice test of quantum complementarity with interfering single photons. *Phys. Rev. Lett.* **100**(22) (2008)

Chapter 2

Direct Measurement of the Photon's Spatial Wave Function



Mohammad Mirhosseini, Jeff S. Lundeen and Robert W. Boyd

Abstract We overview recent progress in the tomography of structured light fields, with an emphasis on the method of direct measurement. Direct measurement provides a scalable and easy- to-implement approach for characterizing the transverse structure of single photons. This protocol is particularly attractive in light of the emerging role of high-dimensional optical states as a resource for encoding quantum information. We present a summary of various implementations of this technique that aim to characterize the spatial degree of freedom of the optical field.

2.1 Introduction

Measurement lies at the heart of quantum mechanics. The significance of measurement, as a theoretical concept, is reflected in a postulate of quantum mechanics, which is entirely devoted to defining the act of measurement. However, determining an arbitrary quantum state of an individual system is prohibited by the quantum no-cloning theorem [1–3]. Quantum information physicists often rely on the method of *quantum-state tomography* to find the quantum state. In this approach, the statistical results of a sequence of projective measurements are combined to infer the state

M. Mirhosseini (✉) · R. W. Boyd

The Institute of Optics, University of Rochester, 275 Hutchison Road, Rochester, NY 14627, USA

e-mail: mohmir@caltech.edu

R. W. Boyd

e-mail: boydrw@mac.com

J. S. Lundeen · R. W. Boyd

Department of Physics, University of Ottawa, 25 Templeton Street, Ottawa, ON K1N 6N5, Canada

e-mail: jlundeen@uottawa.ca

of a quantum system [4–12]. Although well established, quantum-state tomography becomes increasingly challenging for a high-dimensional state due to the number of required experimental measurements and also due to the complexity of the procedure for reconstruction of the quantum state from the experimental data [9, 13]. Recently, an alternative approach has been developed for directly measuring the quantum wave function through the use of weak values [14–19]. A weak value is the result of a weak measurement, which is a form of generalized weak measurement [20]. *Direct measurement* is based on the idea that a weak value, unlike the result of a “conventional” measurement, can be a complex quantity. Hence, the real and imaginary parts of the wave function can be accessed by measuring a complex weak value that is found by a sequence of weak and strong measurements performed in complementary bases, such as the position and momentum bases. We begin this chapter by providing a brief review of direct measurement. We then discuss a variety of experimental implementations of this technique as applied to structured light fields, and conclude the chapter by providing a comparison of direct measurement with quantum-state tomography.

2.2 Strong and Weak Measurements

In this section, we provide a brief review of the mathematical model used to describe conventional measurements often encountered in quantum mechanics (i.e. projective measurements). Subsequently, we will review a generalization of this framework, which we will use for describing weak measurements. The von Neumann interaction is a mathematical model of measurement in which the process of measurement is treated quantum mechanically. In this model, the act of measurement is described as a coupling between a measured system and another quantum system that represents the measurement apparatus, often called a pointer. The Hamiltonian for the interaction of an observable \hat{A} of the quantum system coupled with the momentum \hat{P} of the pointer can be written as

$$\hat{\mathcal{H}} = g\hat{A} \otimes \hat{P}, \quad (2.1)$$

where g is a coupling constant. We have used the tensor product notation \otimes to emphasize that the operators \hat{A} and \hat{P} act on different Hilbert spaces. However, we are going to drop the outer product sign when applying the results of this section later in this chapter. For an interaction time of duration T , the resulting time evolution operator is

$$\hat{U}(T) = \exp\left(\frac{-igT\hat{A} \otimes \hat{P}}{\hbar}\right). \quad (2.2)$$

We choose to work in the basis spanned by the eigenstates of \hat{A} in order to evaluate the effect of \hat{U} on the joint system-pointer quantum state. We note that \hat{A} is diagonal in this basis.

$$\hat{A} = \sum_a |a\rangle a\langle a|, \quad (2.3)$$

where a is the eigenvalue of \hat{A} . We assume an unentangled initial system-pointer quantum state $|\psi(0)\rangle = |I\rangle |\phi\rangle$. Here, $|\phi(x)\rangle$ is the initial state of the pointer. The initial state of the quantum system, $|I\rangle$, can be written as a superposition of \hat{A} eigenstates

$$|I\rangle = \sum_a \alpha_a |a\rangle. \quad (2.4)$$

The effect of the time evolution operator on the joint system-pointer quantum system is thus found as

$$\hat{U}(T)|\psi(0)\rangle = \hat{U}(T) \left(\sum_a \alpha_a |a\rangle \otimes |\phi(x)\rangle \right). \quad (2.5)$$

Note that \hat{P} is the generator of translations of the pointer, and hence

$$e^{-i\frac{x_0}{\hbar}\hat{P}} \phi(x) = e^{-x_0 \frac{d}{dx}} \phi(x) = \phi(x - x_0), \quad (2.6)$$

where $\phi(x) = \langle x|\phi(x)\rangle$. Subsequently, the state of the pointer and the system after the evolution becomes.

$$\hat{U}(T)|\psi(0)\rangle = \sum_a \alpha_a |a\rangle \otimes |\phi(x - \frac{gTa}{\hbar})\rangle. \quad (2.7)$$

It is evident from (2.7) that the position of the pointer has been (coherently) shifted by a value that is proportional to the value a of observable \hat{A} . This is as far as von Neumann's quantum model of measurement can go. It leaves the total system in an entangled state of the system and pointer. We now consider what a conventional measurement of the pointer's position would reveal. We assume an initial Gaussian pointer state at $t = 0$

$$\phi(x) = \left(\frac{1}{2\pi\sigma^2} \right)^{1/4} \exp \left(-\frac{x^2}{4\sigma^2} \right). \quad (2.8)$$

For such a pointer state, the initial position of the pointer is confined to be within a range of positions $\Delta x \propto \sigma$. After the interaction time T , the value a of \hat{A} can be resolved provided that the strength of the measurement is sufficiently large, $gTa \gg \hbar\sigma$. In this regime, each pointer position in (2.7) is correlated with a single value of a . Consequently, the measurement of the pointer's position projects the initial state of the quantum system $|I\rangle$ to $|a\rangle$ with a probability of $|\langle a|I\rangle|^2 = |\alpha_a|^2$.

A well-known example of a projective measurement is the Stern-Gerlach device. Such a device measures the z component of the spin of a spin-1/2 particle by causing

it to interact with a magnetic field along the z direction. In this case the measurement operator is $\hat{A} = \sigma_z$ and the interaction Hamiltonian can be written as

$$\hat{\mathcal{H}} = -g\mu z\sigma_z. \quad (2.9)$$

In this example, the measured system is the spin of the particle and the pointer is simply another degree of freedom, z , of the same particle. Here, $\mu\vec{\sigma}$ is the magnetic moment operator of the particle, and g is the coupling constant which is set by the strength of the magnetic field in the apparatus. Note that in this case the observable $\hat{A} = \sigma_z$ is coupled to the position of the pointer rather to its momentum. Since z is the generator of translations for P_z , the pointer will receive a “kick” in a direction set by its spin. The resulting change of the momentum of the operator can be measured by allowing the beam of particle(s) to freely propagate. In the far field, the centroid position of the beam is correlated with the momentum of the particle inside the Stern-Gerlach apparatus.

We now review the case of weak measurement, closely following the original treatment by Aharonov et al. [20]. In a weak measurement, the strength of the coupling is not sufficient for resolving the different eigenvalues of \hat{A} because the uncertainty in the position of the pointer is larger than than the displacement to be measured, that is $\sigma \gg gTa/\hbar$. In this situation, the joint system-pointer state after the interaction can be approximated by using the first few terms of a Taylor series

$$\begin{aligned} |\psi(T)\rangle &= \exp\left(\frac{-igT\hat{A} \otimes \hat{P}}{\hbar}\right)|I\rangle|\phi\rangle \\ &= |I\rangle|\phi\rangle - \frac{igT}{\hbar}\hat{A}|I\rangle\hat{P}|\phi\rangle + \dots \end{aligned} \quad (2.10)$$

In contrast to conventional measurement, to first order in gT the system and pointer are unentangled. It follows that a measurement of the pointer position will not project the system to $|a\rangle$ but rather leave system unchanged. What if the system is now observed and found to be in final state $|F\rangle$? This process is known as “post-selection”. Conceptually, it is equivalent to limiting our analysis to a sub-ensemble of systems that start with an initial state $|I\rangle$, go through the interaction with the pointer for a period of time T , and then end up in the final state $|F\rangle$. After projecting the state of the evolved system on $|F\rangle$ we find that

$$\begin{aligned} \langle F| \exp\left(\frac{-igT\hat{A} \otimes \hat{P}}{\hbar}\right)|I\rangle|\phi\rangle \\ = \langle F|I\rangle|\phi\rangle - \frac{igT}{\hbar}\langle F|\hat{A}|I\rangle\hat{P}|\phi\rangle + \dots \end{aligned} \quad (2.11)$$

Note that at this stage we have post-selected on the state of the quantum system and we are left with an expression for the state of the pointer. This expression can be normalized by dividing it to $\langle F|I\rangle$, to find

$$|\phi_{\text{FI}}\rangle = |\phi\rangle - \frac{igT}{\hbar} \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \hat{P}|\phi\rangle + \dots \quad . \quad (2.12)$$

The subscript “FI” has been used to emphasize that we are limiting our analysis to the sub-ensemble of systems that are initially prepared in the state $|I\rangle$ and end up in the post-selected state $|F\rangle$. Recall that in the absence of interaction the post-selection procedure succeeds with a probability of $P_{\text{success}} = |\langle F|I\rangle|^2$. This statement remains approximately correct for weak measurements. It is evident from (2.12) that the state of the pointer is modified via the quantity $A_w = \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle}$. This quantity is known as the OAM of the observable \hat{A} . Below, we show that the weak value is observable through the average position and momentum of the pointer. Our analysis closely follows that of Lundeen and Resch [21].

To proceed, we assume the pointer is initially prepared in a Gaussian state described above in (2.8). Although each trial provides little information, one can find the mean position of the pointer in the post-selected sub-ensemble by averaging repeated measurements. To the first order in gT it is

$$\begin{aligned} \langle \hat{X} \rangle_{\text{FI}} &= \langle \phi_{\text{FI}} | \hat{X} | \phi_{\text{FI}} \rangle \\ &= -\frac{igT}{\hbar} \text{Re} \left(\frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right) \langle \phi_{\text{FI}} | (\hat{X}\hat{P} - \hat{P}\hat{X}) | \phi_{\text{FI}} \rangle \\ &\quad + \frac{gT}{\hbar} \text{Im} \left(\frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right) \langle \phi_{\text{FI}} | (\hat{X}\hat{P} + \hat{P}\hat{X}) | \phi_{\text{FI}} \rangle \\ &= gT \text{Re} \left(\frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right) . \end{aligned} \quad (2.13)$$

The last step of the derivation above is performed by using the commutation and anti-commutation relations between position and momentum operators. To first order in gT , the position of the pointer state is shifted by the real part of

$$A_w = \frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle}, \quad (2.14)$$

which is the weak value. The imaginary part of the appears in the conjugate pointer variable, momentum. Specifically, the expectation value of the pointers' momentum is given by

$$\begin{aligned} \langle \hat{P} \rangle_{\text{FI}} &= \langle \phi_{\text{FI}} | \hat{X} | \phi_{\text{FI}} \rangle \\ &= -\frac{igT}{\hbar} \text{Re} \left(\frac{\langle F|\hat{A}|I\rangle}{\langle F|I\rangle} \right) \langle \phi_{\text{FI}} | (\hat{P}^2 - \hat{P}^2) | \phi_{\text{FI}} \rangle \end{aligned}$$

$$\begin{aligned}
& + \frac{gT}{\hbar} \text{Im} \left(\frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right) \langle \phi_{\text{FI}} | (\hat{P}^2 + \hat{P}^2) | \phi_{\text{FI}} \rangle \\
& = \frac{\hbar g T}{2\sigma^2} \text{Im} \left(\frac{\langle F | \hat{A} | I \rangle}{\langle F | I \rangle} \right). \tag{2.15}
\end{aligned}$$

Thus, we have shown that the shift in the expectation values of \hat{X} and \hat{P} of the pointer are proportional to the real and imaginary parts of the weak value.

Because the denominator can be close to zero, a can extend beyond the range of eigenvalues of the measured operator, in contrast to the result of a projective measurement [22]. This property, known as weak-value amplification (WVA), has been used to sensitively measure a variety of effects, such as transverse beam deflections [23–26], phase differences [27], time delays [28], and nonlinear optical effects [29]. In addition to weak-value amplification, weak values were recently used for directly measuring the quantum wave function [14, 16, 30]. We will further discuss the role of complex weak values in measuring the quantum wave function in the following sections.

2.3 Direct Measurement of the Wave Function

We now summarize the direct measurement formalism, following the treatment by Lundeen and co-workers [14]. We consider a weak measurement of the system's position via a projector at point x , i.e. $\hat{A} = |x\rangle\langle x|$. This is followed by post-selection on an eigenstate of momentum, $|p\rangle$. The weak value for this measurement can be written as

$$\langle \pi_x \rangle_w = \frac{\langle p | x \rangle \langle x | \psi \rangle}{\langle p | \psi \rangle}. \tag{2.16}$$

Substituting the Fourier transform relation between x and p , $\langle p | x \rangle = \frac{1}{\sqrt{2\pi}} e^{ipx/\hbar}$, we obtain

$$\langle \pi_x \rangle_w = \frac{e^{ipx/\hbar} \psi(x)}{\sqrt{2\pi} \phi(p)}. \tag{2.17}$$

Here, $\psi(x) = \langle x | \psi \rangle$ and $\phi(p) = \langle p | \psi \rangle$ are the wave function representations in the position and momentum bases, respectively. This expression can be further simplified by considering the special case of post-selection on $|p = 0\rangle$

$$\langle \pi_x \rangle_w = k \psi(x). \tag{2.18}$$

It is evident that the weak value is directly proportional to the quantum wave function. The proportionality constant, $k = \frac{1}{\sqrt{2\pi} \phi(0)}$, can be eliminated by impos-

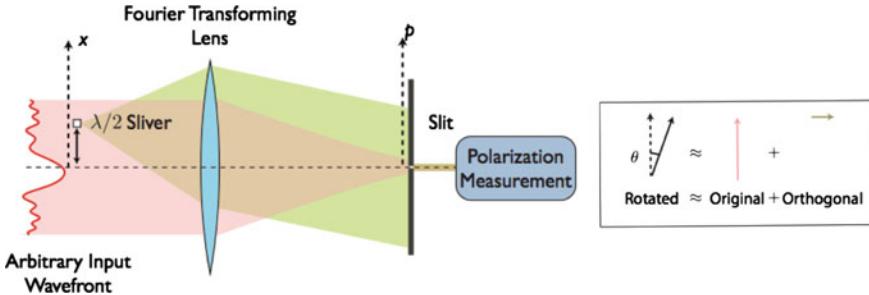


Fig. 2.1 Left panel: experimental setup for direct measurement of the wavefunction. Right panel: a small polarization rotation can be pictured as the superposition of the original polarization plus a small orthogonally polarized component

ing the normalization condition of the wave function, or by simply measuring $\text{Prob}(p = 0)$.

Figure 2.1 shows a schematic setup for the experiment in [14]. In this realization, instead of using the position of the photon as a pointer we use its polarization. Accordingly, instead of a linear shift of the pointer one would expect a rotation, and so, it behaves like a pointer on a circular dial. The wave function $\psi(x)$ is the transverse structure of the optical field. Initially, the beam of photons is prepared to have a vertical polarization. To perform a weak measurement, the polarization of the optical field is rotated by a small angle, α , at the point x by placing a small half-wave plate at that location. The strong measurement of momentum is performed by a slit located after a Fourier-transforming lens. In analogy to the position shift seen in Sect. 2.2, one would expect the real part of the weak value to be proportional to the rotational shift of the pointer around the dial. This can be found by measuring the imbalance between the probability to detect a diagonal and anti-diagonal polarized photon. Formally, this imbalance is given by $\langle s | \sigma_x | s \rangle$, where $|s\rangle$ denotes the polarization state of the photon passing through the slit and σ_j is the $j = x, y, z$ Pauli matrix. The vertical polarization is the eigenstate of σ_z with a -1 eigenvalue. (Note, this differs from the convention that is used in atomic physics, for instance.) In analogy to the pointer momentum shift in Sect. 2.2, the imaginary part of the weak value would then be proportional to the average value of the conjugate variable, which is σ_y . This average is the imbalance between the probabilities of the photon having a right-hand circular and left-hand circular polarization. Putting the real and imaginary parts together, one can show that the weak value is given by Lundein et al. [14]

$$\langle \pi_x \rangle_w = \frac{1}{\sin \alpha} (\langle s | \sigma_x | s \rangle - i \langle s | \sigma_y | s \rangle). \quad (2.19)$$

Classically, the expectation values of σ_x and σ_y are equivalent to the S_2 and S_3 polarization Stokes parameters normalized to S_0 .

Reference [14] proposes and demonstrates the concept of direct measurements. The authors have used direct measurement to characterize the transverse wave func-

tion of single photons exiting a single mode fiber. By scanning the small half-wave plate sliver, it is possible to measure the real and imaginary parts of the field at each point and then convert the results to amplitude and phase. The authors have repeated the experiment with a classical beam of light and have verified that the results remain identical to those from the experiment with single photons [31]. They have also repeated the experiment, replacing the slit with a camera, thereby keeping all the outcomes of the strong measurement [32, 33]. Finally, they also demonstrated that a similar method can be used to directly measure the density matrix [31, 34]. While a complete comparison of the direct measurement with quantum tomography remains to be drawn, the simplicity of its implementation has provided ample motivation for different experimental realizations [15, 16, 30, 35].

2.4 Direct Measurement of the OAM State Vector

The original implementation of direct measurement considered position and momentum as the two conjugate variables required for describing a quantum state. However, direct measurement can be applied to any pair of conjugate quantum observables that describe the quantum system under study. In this section, we discuss the measurement of the transverse state vector of a photon in the orbital angular momentum (OAM) basis following the original development by Malik et al. [16]. We begin by expressing the state of a photon as a superposition of states in the OAM basis as

$$|\Psi\rangle = \sum_{\ell} a_{\ell} |\ell\rangle. \quad (2.20)$$

To find the OAM state vector, we need to find the complex probability amplitudes a_{ℓ} . Following the direct measurement formalism, we consider a pair of measurements in the orbital angular momentum and the angular position bases. In analogy to linear momentum and position, the pair of OAM and angular position variables are connected via a discrete Fourier transformation [36–38]. We define an angular state as

$$|\theta\rangle = \frac{1}{\sqrt{d}} \sum_{\ell=-N}^{\ell=+N} \exp\left(-\frac{2\pi i}{d}\theta\ell\right) |\ell\rangle. \quad (2.21)$$

Where we have chosen to work in a finite-dimensional state space spanned by the orbital-angular-momentum eigenvectors $|\ell\rangle$ with $\{|\ell| \leq N\}$ and $d = 2N + 1$. Using the definition above, it is straightforward to verify that the quantity $c = \langle\theta_0|\ell\rangle/\langle\theta_0|\Psi\rangle$ is a constant with respect to $|\ell\rangle$ for the case where $\theta_0 = 0$. We now multiply both sides of (2.20) by the constant c to find

$$c|\Psi\rangle = c \sum_{\ell} |\ell\rangle \langle\ell|\Psi\rangle = \sum_{\ell} |\ell\rangle \frac{\langle\theta_0|\ell\rangle \langle\ell|\Psi\rangle}{\langle\theta_0|\Psi\rangle} = \sum_{\ell} \langle\pi_{\ell}\rangle_w |\ell\rangle. \quad (2.22)$$

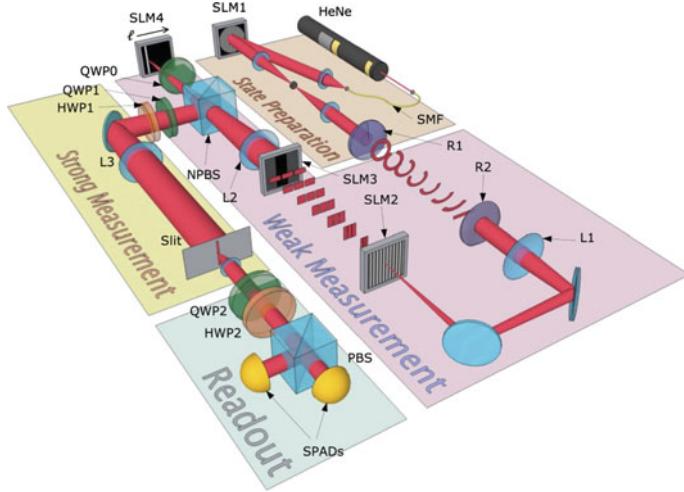


Fig. 2.2 Configuration of the experiment for direct measurement of a 27-dimensional OAM state vector, as described by Malik et al. [16]. The conjugate measurements of OAM and angular position in this experiment are enabled by using an OAM mode sorter described in [39]

In this expression, we have introduced the quantity $\langle \pi_\ell \rangle_w$, which is proportional to the probability amplitude a_ℓ from (2.20). This quantity is in fact the weak value for the OAM projection operator, and using the results from the previous section we can conclude that it is equal to the average result obtained by making a weak projection in the OAM basis ($\hat{\pi}_\ell = |\ell\rangle\langle\ell|$) followed by a strong measurement in the conjugate basis of angular position (θ). Thus, it is established that the scaled complex probability amplitudes ca_ℓ can be directly obtained by measuring the OAM weak value $\langle \pi_\ell \rangle_w$. Following the measurement procedure, we can normalize the state $|\Psi\rangle$ to eliminate the constant c .

Malik et al. have utilized the polarization of photons as the measurement pointer for realizing direct measurement of the OAM state vector [16]. Figure 2.2 shows the schematic diagram of the setup used in their experiment. In this work, the conjugate pair of OAM and angular position are accessed by using an OAM mode sorter, which maps the polar coordinates to Cartesian coordinates [39]. The weak measurement is performed by rotating the polarization of an OAM component of the input beam, and it is followed by passing the beam through a slit that performs a strong measurement of the angular position. The OAM weak value is subsequently read out by measuring the average change in the photon's linear and circular polarization.

The authors in [16] have used direct measurement to characterize photons with a $\psi(x)$ in shape of an angular wedge. An angular wedge can be theoretically described as a superposition of OAM modes with the probability amplitudes,

$$a_\ell = k \operatorname{sinc}\left(\frac{\Delta\theta\ell}{2}\right). \quad (2.23)$$

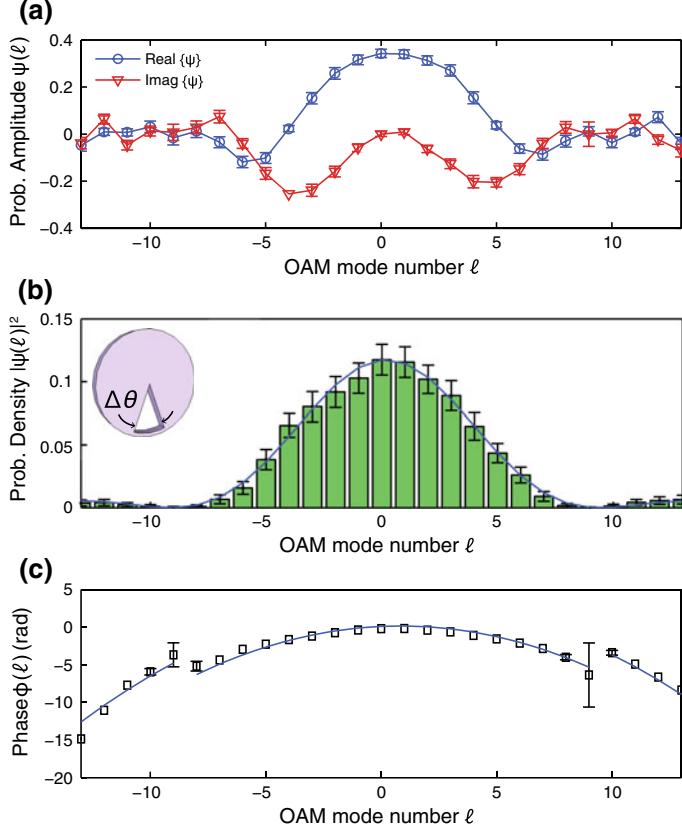


Fig. 2.3 Experimental data showing direct measurement of a 27-dimensional state vector in the OAM basis by by Malik et al. [16]. The photons are prepared in form of an angular wedge (inset of (b)) with the angular width of $\Delta\theta = 2\pi/9$ rad. Panel a shows the measured real (blue circles) and imaginary parts (red triangles) of the state vector. The probability density $|\Psi(\ell)|^2$ (b), and phase $\phi(\ell)$ (c) are calculated using the real and imaginary parts shown in (a)

Figure 2.3 demonstrates the measured probability amplitude for this state. It is evident that the probability amplitudes are in close agreement with the theory fit. Interestingly, the data manifests the expected sign change of the probability amplitude in the vicinity of the nulls in the sinc function [(note the phase shifts of π in panel (c)]. The authors attribute the quadratic shape of the measured phase to an artifact in the experimental apparatus. The solid lines provide fits to the theory after taking into accounts the experimental misalignments.

As another test for the reliability of the measurements, Malik et al. have analyzed the effect of rotation on a photon carrying a broad range of angular momenta. Considering that the orbital angular momentum operator is the generator of rotations, the rotation of a state results in a phase shift among its OAM components. A rotation by

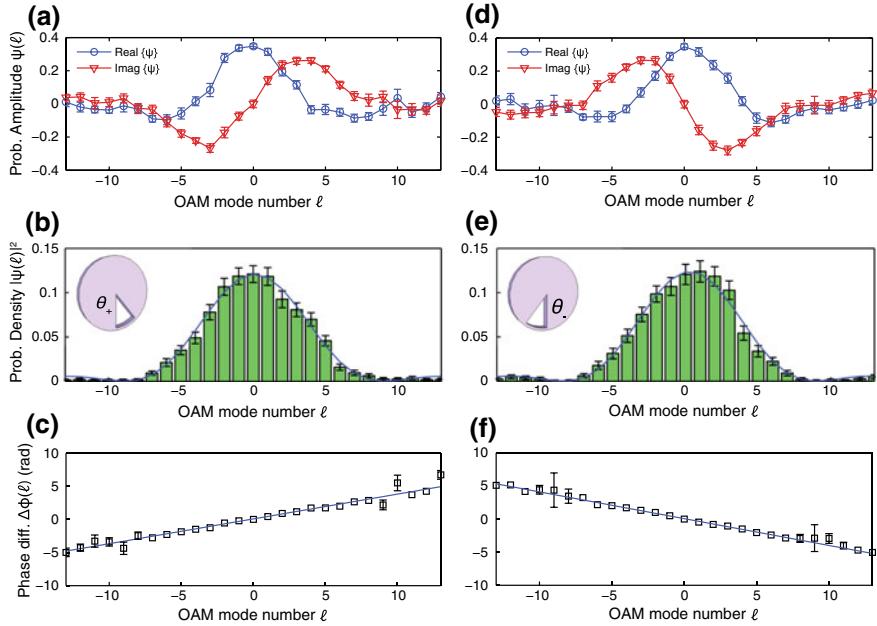


Fig. 2.4 Experimental data showing the direct measurement of a rotated high-dimensional state vector by by Malik et al. [16]. The angular wedge state is rotated by angles $\theta_\pm = \pm\pi/9$ rad [insets of (b) and (e)]. Panels **a** and **d** present the measured real and imaginary parts of the rotated state vectors. Panels **b** and **e** show the probability densities calculated using the real and imaginary parts. Panel **c** and **f** present the difference between the calculated phase for each case and the phase of the unrotated case from Fig. 2.3c

an angle θ_0 can be expressed by the unitary operator $\hat{U} = \exp(i\hat{L}_z\theta_0)$, where \hat{L}_z is the angular momentum operator. Applying this operator to our quantum state $|\Psi\rangle$ with \hat{U} , we get

$$|\Psi'\rangle = \hat{U}|\Psi\rangle = \sum_{\ell} k \operatorname{sinc}\left(\frac{\Delta\theta\ell}{2}\right) e^{i\ell\theta_0} |\ell\rangle. \quad (2.24)$$

Thus, rotation by an angle θ_0 manifests as an ℓ -dependent phase $e^{i\ell\theta_0}$ in the OAM basis. Figure 2.4 presents the measured probability amplitudes for a rotated angular wedge. It is evident that rotation of the state results in a relative phase among the OAM components that scales linearly with the OAM index. Further, the authors have shown that it is possible to reliably infer the rotation angle θ_0 by finding the slope of the linear phase structure [16].

This experiment clearly shows the possibility of implementing direct measurements for various degrees of freedom that describe the wave function. Direct measurement provides an easy-to-implement recipe for characterizing the state with no requirements for post-processing, provided one can access two conjugate variables for performing weak and strong measurements. The results clearly demonstrate the

fundamental relation between phase and rotation in the OAM basis. Further, the ability to characterize the OAM state vector may lead to significant applications in quantum communications considering the emerging role of twisted photons as carriers of quantum information [40].

2.5 Compressive Direct Measurement

Direct measurement provides a simple way to find the quantum wave function with no post-processing. However, the measurement of the entire wave function, $\psi(x)$, in a domain such as position requires repeating the experiment for all the possible value of the position variable x . This *scanning* procedure becomes increasingly difficult when the number of discrete values for x increases and more significantly so, when the state is defined for multiple variables such the pair of coordinate variables x and y which describe the transverse plane of an optical beam. In this section, we review *compressive direct measurements* (CDM) introduced by Mirhosseini et al. [30]. The CDM method combines the benefits of direct measurement with a novel computational technique known as compressive sensing [41–46]. It has been shown that by utilizing this approach, the wave function of a high-dimensional state can be estimated with a high fidelity using much fewer number of measurements than the standard direct measurement approach.

It was previously shown by Lundeen et al. that the weak measurement of the position projection operator followed by a strong measurement of momentum leads to the direct measurement of the wave function [14]. In the CDM approach, the position projector operator is replaced with a more general measurement operator known as a *sensing operator*, which is formed as a general combination of projection operators for multiple positions. Consider a weak measurement of the sensing operator $\hat{Q}_m = \sum_j Q_{m,j} \hat{\pi}_j$, which is a linear combination of the position projectors $\hat{\pi}_j = |x_j\rangle\langle x_j|$ weighted with coefficients $Q_{m,j} \in \mathbb{R}$. Mirhosseini et al. have shown that the weak measurement of the sensing operator defined above followed by a post-selection on momentum results in a weak value that is proportional to a linear combination of the wave function at multiple points [30]. This can be seen by calculating the expectation values for the Pauli operator of the pointer after post-selection

$$\begin{aligned}\bar{\sigma}_{x,m} &\equiv \langle s_m | \hat{\sigma}_x | s_m \rangle = \kappa \sum_j Q_{m,j} \operatorname{Re}[\psi_j], \\ \bar{\sigma}_{y,m} &\equiv \langle s_m | \hat{\sigma}_y | s_m \rangle = -\kappa \sum_j Q_{m,j} \operatorname{Im}[\psi_j].\end{aligned}\quad (2.25)$$

Here, $\hat{\sigma}_x = |H\rangle\langle V| + |V\rangle\langle H|$, and κ is a constant number quantifying the strength of the weak measurement [30]. In the above relations, $\operatorname{Re}[\psi_j]$ and $\operatorname{Im}[\psi_j]$ are the real and the imaginary parts of $\psi(x_j)$, respectively. We have assumed that the pointer is initially prepared in the horizontal state $|H\rangle$.

Equations (2.25) show that for a sensing matrix, \hat{Q}_m , the measurement result is proportional to a linear combination of the real and imaginary part of the wave function at different positions x_j weighted with coefficients $Q_{m,j}$. Denoting the measurement results as $\phi_m = \frac{1}{\kappa}[\bar{\sigma}_{x,m} - i\bar{\sigma}_{y,m}]$, the experiment can be repeated with multiple different sensing operators to obtain a system of linear equations, $\phi = \mathbf{Q} \psi$. Here, the wave function ψ is treated as a vector of unknown variables and ψ is a vector that summarizes all the measurement results. In the expanded form the system of linear equations read

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_M \end{pmatrix} = \begin{pmatrix} Q_{1,1} & Q_{1,2} & \cdots & Q_{1,N} \\ Q_{2,1} & Q_{2,2} & \cdots & Q_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{M,1} & Q_{M,2} & \cdots & Q_{M,N} \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix}. \quad (2.26)$$

Here, $m \in \{1 : M\}$ and $n \in \{1 : N\}$, where M is the total number of sensing operators and N is the dimension of the wave vector ψ .

The wave function ψ can be found in the case in which the linear system of equations above have a unique solution. This happens for the special case $M = N$, provided that the measurement matrix \mathbf{Q} is not singular. However, a more interesting case happens for $M \leq N$, where we intend to find the unknown N -dimensional vector ψ with fewer than N measurement results. Obviously, a unique solution cannot be found in this case because of the lack of complete information about the wave function. However, it may still be possible to infer the wave function provided there exists some *prior* information about the quantum state under investigation. This can, in fact, be achieved by using the idea of compressive sensing, where this prior information is expressed in terms of sparsity. If the wave function under investigation (ψ) can be mapped to a sparse vector (i.e. a vector in which most of the elements are zero) by using a known linear transformation (\mathbf{T}), it can be retrieved from the measurement results by solving a convex optimization problem [47]

$$\min_{\psi'} ||\mathbf{T}\psi'||_{\ell_1}, \text{ subject to } \mathbf{Q}\psi' = \phi, \quad (2.27)$$

where $|| \cdot ||_{\ell_1}$ represents the 1-norm. If the number of experimental results, M , satisfy the inequality, $M > O[k \log N]$, the wave function ψ can be retrieved with an overwhelming probability[48]. Here, k is the number of nonzero elements of $\mathbf{T}\psi$.

Vectors with spatial correlations between neighboring elements (such as an image, for example) are likely to be sparse in discrete cosine transform, wavelet transform, and the discrete gradient domains [47, 49]. Mirhosseini et al. have used the discrete gradient operator for recovering the transverse structure of photons in a laser beam. Figure 2.5 shows the schematic of their experiment. A spatial light modulator (SLM) is used along with a pair of quarter-wave plates to control the polarization of the beam in a pixel-by-pixel fashion. The authors realize a binary sensing matrix by either rotating the polarization of the incoming beam at each pixel or by leaving it intact. Following the weak measurement, the post-selection is performed by using

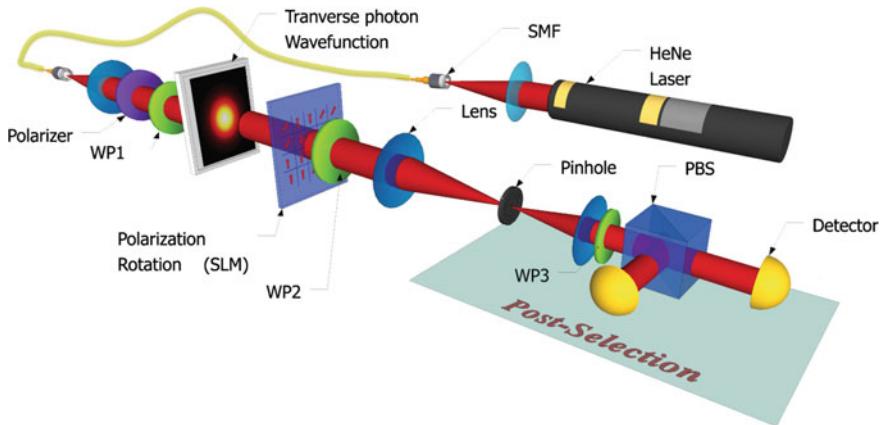


Fig. 2.5 The schematic of the experimental setup for compressive direct measurement described by Mirhosseini et al. [30]. A liquid crystal spatial light modulator has been used along with a pair of quarter-wave plates to rotate the polarization of the incoming laser beam

a lens a pinhole. In the final stage, the polarization of photons from the pinhole are characterized to find the weak value corresponding to the measurement matrix. The experiment is repeated after the sensing matrix is updated by digitally controlling the SLM.

The left panel of Fig. 2.6 shows the retrieved wave function from the experimental results from [30]. The left column presents data from a pixel-by-pixel scan of the state for $N = 192$. The middle column shows the reconstructed wavefront for for $N = 192$, and $M/N = 20\%$ of total measurements from the CDM method. It is evident that the wave function reconstructed from the 20% of the measurement captures the main features of state. The quality of the reconstructed state can be further enhanced by sampling the wave function at a larger number of pixels. Increasing the number of pixels results in a more smooth function and subsequently increased sparsity in the discrete gradient domain. The right column demonstrates reconstruction $N = 19,200$, and $M/N = 20\%$ of total measurements.

The authors have provided a quantitative comparison of the state found from compressive and the standard direct measurement [30]. The right panel of Fig. 2.6 shows the fidelity of the reconstructed state as a function of the percentage of the measurements ($100 \times \frac{M}{N}$). Fidelity is closeness metric, and is defined for a pair of wave functions as $F(|\psi'\rangle, |\psi\rangle) = |\langle\psi'|\psi\rangle|$. The blue curve shows the fidelity of the state reconstructed with the CDM method, whereas the red curve represents the average fidelity of state reconstructed from linear reconstruction (pseudo-inverse) using the data from a partial pixel-by-pixel measurement of M randomly chosen points. It is evident from the figure that the compressive method provides a fairly accurate reconstruction using as low as 20% of all the measurement. Nonetheless, finding the wave function with a fidelity of near to 1 requires performing all N measurement regardless of the reconstruction method. The nonlinear behavior of fidelity as a function of

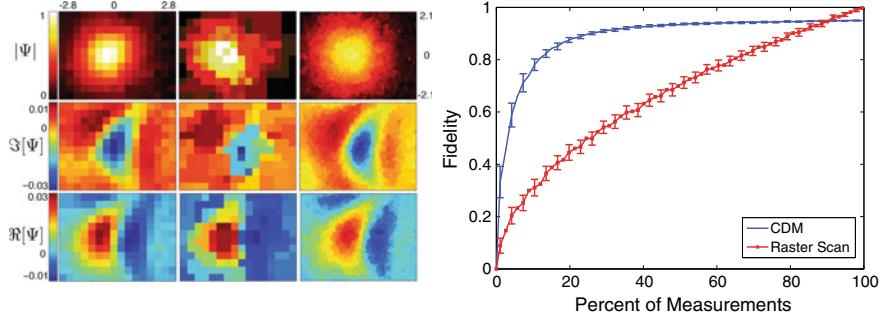


Fig. 2.6 Left panel: experimental data comparing the conventional direct measurement with compressive direct measurement by Mirhosseini et al. [30]. The left column shows the wave function found by a pixel-by-pixel direct measurement ($N = 192$). The middle column shows the reconstructed wave function from compressive direct measurement ($N = 192$, and $M/N = 20\%$). The right column shows reconstructed wave function from compressive direct measurement when the number of pixels is increased ($N = 19,200$ and $M/N = 20\%$). Right panel: the fidelity of the reconstructed state with the target wave function as a function of the percentage of the total measurements from [30]. The fidelity of the state reconstructed with CDM is shown in blue. The fidelity of the state reconstructed from a partial pixel-by-pixel scan with the same number of measurements is shown in red for comparison

the percentage of the measurements is a signature of compressive imaging and has been observed elsewhere previously [49].

The CDM approach allows for finding an unknown wave function from a fraction of the measurements required in the standard direct measurement. This property leads to a significant speed-up that can crucially benefit the tomography of large-dimensional states. In addition to this property, the randomness of the measurement operators used in this approach results in a weak measurement on approximately half of all the photons in the beam for every setting of the experiment. Due to this, the change in the state of the pointer (for example the polarization of the beam after the pinhole in Fig. 2.5) is much more pronounced as compared to the standard direct measurements where only one position eigenstate would be perturbed. This property results in a pronounced signal-to-noise ratio (SNR) in the displacements of the expectation values for the Pauli operators of the qubit pointer. The enhanced SNR subsequently leads to an additional decrease in the measurement time. The authors in [30] have found a 350-fold speed-up in the procedure of characterizing a massive $N = 19,200$ -dimensional state (See Fig. 2.7).

This experiment clearly demonstrates the benefits of combining direct measurements with various numerical algorithms that can speed-up the reconstruction of the wave function. The favorable scaling provided by compressive direct measurement allows for characterizing massive high-dimensional states with only a fraction of measurements required by the standard direct measurement. Besides quantum-state tomography, compressive direct measurement can be used for retrieving phase

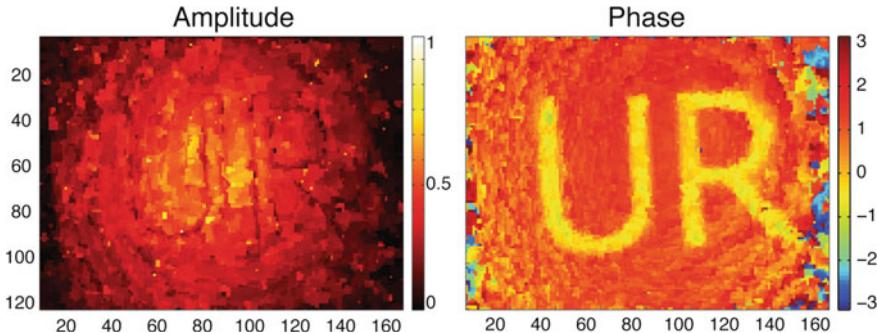


Fig. 2.7 Compressive reconstruction of amplitude and phase for a high-dimensional state by Mirhosseini et al. [30]. The state is created by illuminating a phase mask depicting letters U and R with a phase jump of $\pi/2$. The data has been reconstructed with $N = 19,200$, and $M/N = 20\%$

information in classical applications such as wave-front sensing of THz beams. The exclusive use of single-pixel detectors in this approach provides a simple alternative to using costly and inefficient arrayed detectors in such applications.

2.6 “Scan-Free” Direct Measurement

Direct measurement significantly simplifies the post-processing required for characterizing a quantum state. Nevertheless, in its original form, direct measurement suffers from two major limitation: the need for scanning and the post-selection loss. Direct measurement requires weakly measuring the projection operator for the bases of the Hilbert space of interest in a sequence. Scanning through this sequence becomes increasingly time consuming as the dimension size of the Hilbert space grows. The second limitation arises from discarding a significant number of photons in the post-selection procedure. Since the photons disposed in post-selection generally carry information about the system [50], their discarding further prolongs the measurement procedure.

Lundeen and co-workers have recently extended direct measurement to find the Dirac distribution by performing parallel post-selections [51]. Their scheme is based on using a detector array and thus it avoids post-selection loss. Although the Dirac distribution can be used to find the wave function of a pure state, it generally contains more information since it can also characterize mixed states. Consequently, measurement of the Dirac distribution still requires scanning the weak measurement operators. Shi et al. have devised a technique that substantially simplifies this procedure by interchanging the measurement of position and momentum eigenstates [52]. As detailed below, this approach enables the measurement of the entire wave function by using a detector array and with only one setting of the experimental apparatus.

In scan-free direct measurement, the wave function of a photon in the position is measured weakly by the momentum projection operator ($\hat{\pi}_{p_0} = |p_0\rangle\langle p_0|$), followed by a strong measurement of the position state $|x\rangle$. For this sequence of measurements, we can apply the weak value formula to find

$$\langle \pi_p \rangle_x^w = \frac{\langle x|p_0\rangle\langle p_0|x \rangle}{\langle x|\psi \rangle} = \frac{e^{-ip_0x/\hbar}\tilde{\psi}(p_0)}{\sqrt{2\pi}\psi(x)}. \quad (2.28)$$

here, $\tilde{\psi}(p)$ and $\psi(x)$ denote the wave function in the momentum and position bases, respectively. This expression can be further simplified by considering the special case of weakly measuring the zero-momentum state, i.e. $p_0 = 0$. In this case, we have

$$\langle \pi_p \rangle_x^w = \frac{\nu}{\psi(x)}, \quad (2.29)$$

where $\nu = \tilde{\psi}(0)/\sqrt{2\pi}$ is a constant. It is evident from this relation that the weak value is proportional to the reciprocal of the wave function. Note that finding $\langle \pi_p \rangle_x^w$ for different values of x requires only changing the post-selection state $|x\rangle$ with no need for changing the operator that is weakly measured. Although this interchange of x and p measurements might appear unimportant, this procedure enables “one-shot” acquiring of the entire wave function since parallel strong measurements of position can be achieved by using a detector array. The value of the constant ν does not require any additional measurement since it can be determined by normalizing the measured wave function.

Although direct measurement was originally developed for finding the wave function of a single photon, it can also be employed to characterize the photons in a laser beam, where the transverse wave function can be alternatively understood as a complex optical mode function [31]. Shi et al. have realized the measurement of such a complex optical field using the approach detailed above [52]. Figure 2.8 depicts the schematics of their experimental setup. In this experiment, a 4-f imaging system provides access to the momentum and the position of the photons in the beam. The weak measurement of the $p = 0$ state is performed by digitally controlling the retardance of the pixels at the center of a spatial light modulator, and the subsequent strong measurements of position are done in parallel by using a CCD camera.

The authors generate and characterize vortex modes in the first part of the experiment (See Fig. 2.9). As expected, the weak value gets large close to the center of a vortex modes where the phase is singular and the optical amplitude is zero. The number of pixels in the measured wave function can be found as approximately 1.2 million. This number is found simply by dividing the area of the laser beam (diameter 7 mm) by the area of an individual pixel on the CCD camera ($5.4 \mu\text{m}^2$). The authors have calculated a fidelity of 0.93 for characterization of an $\ell = 3$ OAM modes. Since the weak values proportional to the reciprocal of the wave function amplitude in this approach, accurate measurement becomes a challenge when $|\psi(x)|^2$

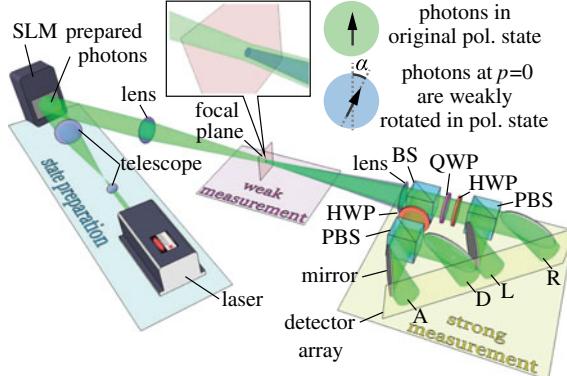


Fig. 2.8 The experimental setup for “scan-free” direct measurement described by Shi et al. [52]. The spatial structure of the beam is controlled by digitally adjusting the parameters of a phase-only grating which is displayed on a spatial light modulator (SLM). The weak measurement of momentum is achieved by using a second SLM (not shown) at the Fourier plane of the first SLM. The polarization rotation requires using a pair of quarter-wave plates placed before and after the second SLM

Fig. 2.9 Real (a) and imaginary (b) parts of measured weak value along with the corresponding wave function (c, d) for a vortex OAM modes with $\ell = 3$ (figure is from [52]). The bottom row shows the measured phase for a number of OAM modes with different indices

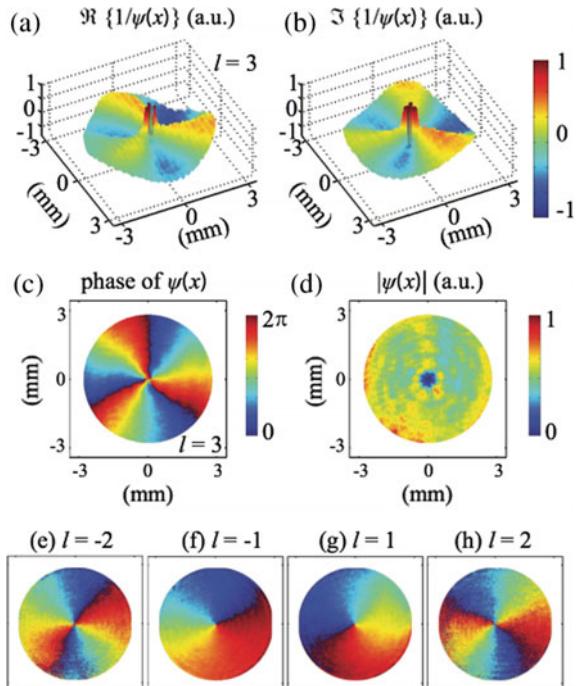
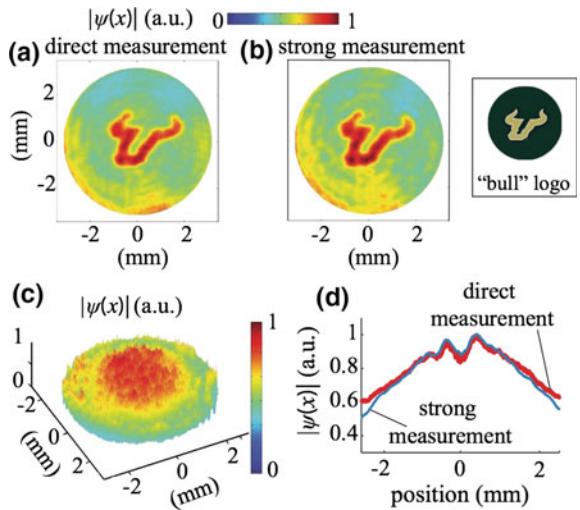


Fig. 2.10 Comparison of direct measurement (a) and strong measurement (b) of an arbitrary shaped beam by Shi et al. [52]. Panel c shows the amplitude of a Gaussian beam from direct measurement. Panel d compares the results of direct measurement and strong measurement for a cross section of the amplitude



becomes small. As another test of the reliability of this technique, Shi et al. have measured an arbitrary shaped beam and compared its amplitude to the result of strong measurements (see Fig. 2.10).

In summary, Shi et al. have realized a variant of direct measurement that is capable of finding the entire wave function with a single setting of the experimental setup and thus eliminates the need for scanning. Although this method was demonstrated by characterizing a strong laser beam, the theoretical approach and experimental setup can both be applied to measuring single photons provided the availability of high resolution single-photon detector arrays. In addition to applications in quantum-state tomography, this technique can potentially benefit classical applications such as astronomy and biological imaging by providing a speedy and easy-to-implement wave-front sensing apparatus.

2.7 Comparison to Quantum State Tomography

The directness of the quantum state determination methods described in this chapter becomes apparent when they are contrasted with quantum-state tomography (QST) [11, 12]. In QST, one begins with an ensemble of identically prepared systems, each described by the quantum state one wishes to determine. This ensemble is divided into sub-ensembles. One performs a measurement in a different basis on each sub-ensemble. Thus, together there is a set of bases. For example, for a polarization state, this set of bases could be the eigenstates of the σ_x , σ_y , and σ_z operators. For the subject of this chapter, the transverse position state $\psi(x)$, this set of bases could be the eigenstates of the operator $x(\theta) = x \cos \theta + p \sin \theta$, for a set of mixing angles θ .

ranging from 0 to 2π (see [53] for an experimental example of this). The frequency of each and every measurement outcome is a data set that one fits a quantum state to. Consequently, the determined quantum state is the one that best predicts the full data set found from measurements in many bases.

In the above description one can see the main three factors contributing to the indirectness of QST. They are: (1) Typically the quantum state is fit to the data using a complicated computation, such as maximum likelihood estimation. (2) The fit is global. That is, the entire state is determined at once. For example, even if one is only interested in the value of the state at a particular point one must nonetheless determine the entire wave function. (3) A diverse set of measurements must be performed on the ensemble. We will discuss these factors in more detail now.

Much work has been done to minimize the size of the ensemble and the number of different measurements required for QST. In all the schemes, the number of bases in the set is of order d and the total number of measurement outcomes across the bases is at least d^2 . Contrast that with the direct measurement methods described in this chapter, which only require two bases. If these bases are mutually unbiased [54] then the measured signal directly gives real and imaginary amplitudes of the state. However, even if the bases are not mutually unbiased, it is still possible to measure the state locally. The measured signals will just need to be corrected by some factor [55]. Thus, the required set of measurements are dramatically reduced for direct measurement methods.

There is a previously existing method that also only requires measurements in two bases. The average result of a simultaneous measurement of x and p give the Husimi Q function, as long as the disturbance from both measurements are balanced [56–58]. The Q function is an $x - p$ two dimensional distribution that is rigorously equivalent to the quantum state, so this is another method of state determination. The direct methods discussed in this chapter are actually closely related to Q function measurement [32]. They both measure a quasi-probability distribution. The definition of these two distributions, the Q function and Dirac distribution, differ only in the ordering of quantum operators. Also, whereas the Q function is found from simultaneous measurements of x and p , the Dirac distribution is found from sequential measurements. However, direct measurement based on weak measurement has the advantage that one can choose the basis in which to determine the state. The Q function is defined in terms of a particular incomplete basis, the coherent state basis. This basis makes it difficult to unambiguously define the Q function for discrete Hilbert Spaces, such as OAM [59, 60].

We now discuss the efficiency of the direct measurement method in comparison to standard QST. The fidelity of the determined state to the actual state fundamentally depends on the size of the ensemble of identical states used. The efficiency of a state determination method is defined by how quickly the fidelity improves as the ensemble size is increased. Consider the direct measurement method: typically only a single post-selection outcome (e.g. $p = 0$) must be kept. The rest of the ensemble is simply discarded and, thus, do not contribute to improving the fidelity. This would seem to be of great detriment to the efficiency of direct measurement. However, while this is the simplest way to perform a direct measurement, it is not the

only way. Another possibility would be to keep and distinguish between all the p outcomes. This strategy has been demonstrated experimentally and gives the Dirac distribution [32, 51], which can also determine mixed quantum states. The scan-free and compressive sensing methods described in Sects. 2.5 and 2.6 of this chapter also address the post-selection drawback.

Even ignoring the post-selection, the direct measurement method is at an apparent disadvantage to QST. Essentially, weak measurement trades measurement disturbance for measurement certainty. That is, the result of a measurement on one member of the ensemble is necessarily designed to be ambiguous and, thus, the information gain will be low. This suggests that the fidelity will improve slowly as the ensemble size is increased and hence the efficiency will be low. Two papers have looked at fidelity of direct measurement quantitatively [61, 62]. Their conclusions support these heuristic arguments: QST is more efficient than direct measurement. However, one of the papers suggests a simple modification of the direct measurement procedure, consisting of additional measurements, that allows one to increase the strength of the measurement coupling. In turn, this increases the information gained per ensemble member. Indeed, the efficiency of this modified direct measurement approaches that of QST [62].

Both QST and direct measurement have many other possible variants that can increase their efficiency. In QST, it is known that the choice of the set of measurement bases can affect efficiency. Improved efficiencies are attained for sets of mutually unbiased bases or a Symmetric Informationally Complete POVM. In contrast, the effect of one's choice of bases is relatively unexplored in the direct measurement method. Another variant exploits prior information about the state to increase efficiency. Compressive sensing does this and is used both in direct measurement (see Sect. 2.5) and QST [5, 44]. Another variant, adaptive measurement, is known to achieve optimal efficiency in QST [63, 64]. In this method, one changes the measurement bases depending on the outcome of measurement conducted on the previous ensemble member. Finally, there are likely many other ways, unrelated to weak measurement, to directly measure state amplitudes [65], although they may not be as experimentally straightforward.

2.8 Summary

Direct measurement has a number of practical benefits. The experimental procedure is relatively straightforward. Put simply, the quantum state can be measured by weakly measuring one variable followed by a strong measurement of a complementary variable. Moreover, the real and imaginary components of the wave function, or more generally, quantum state amplitudes, appear directly on the measurement apparatus. Consequently, direct measurement sidesteps elaborate experimental procedures and arduous computational quantum-state-reconstruction algorithms. These features make direct measurement an attractive option for measuring high dimensional states, such as those that describe a photon's spatial wave function. As exam-

ples, we described how direct measurement can be used to measure orbital angular momentum states and the transverse spatial wave function in one dimension. Other areas of physics have large dimensional states at their foundation, including quantum computing and quantum biology. Direct measurement may prove useful for characterizing these states as well.

Since the original proposal, the idea for direct measurement of quantum states has been extended and improved in a variety of ways. One variation that we described eliminates the need to scan the position of the first measurement. It does so by retaining all the outcomes of the second measurement. If one retains these outcomes and then also scans the first measurement then this determines mixed quantum states by directly giving the Dirac distribution of the system. Improving the efficiency of quantum state determination methods is an important research focus. We described the method of compressive direct measurement and reviewed experimental evidence that it improves the efficiency compared to the original proposal. Other recent variations have demonstrated efficiency improvements [62, 65] and we expect that more will come.

As important as its potential practical use for determining quantum states, direct measurement also has the potential to change our understanding of quantum physics. Quantum states are often thought of as an abstract device for calculating measurement results. Even experts in state determination often see them as shadowy entities: “Quantum states may comprise complementary features that cannot be measured simultaneously and precisely. Consequently, we can not see quantum states directly...” wrote Leonhardt in his textbook, *Measuring the Quantum State of Light* [66]. Seeing quantum states directly is of fundamental importance since it gives them a precise and grounded meaning in terms of a measurement procedure. i.e. an operational definition. However, one should not misunderstand this operational meaning to imply that the quantum state is “real” in the sense used in Bell’s inequality. Rather, an operational definition avoids the deep rooted conceptual assumptions that may limit our understanding of nature. Consequently, at both a practical and foundational level direct measurement is a promising new concept.

MM and RWB acknowledge support by the US Office of Naval Research. JL and RWB acknowledge support by the Canada Research Chairs (CRC) Program, the Canada First Research Excellence Fund (CFREF), and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

1. W.K. Wootters, W.H. Zurek, A single quantum cannot be cloned. *Nature* **299**(5886), 802 (1982)
2. D. Dieks, Communication by EPR devices. *Phys. Lett. A* **92**(6), 271 (1982)
3. P.W. Milonni, M.L. Hardies, Photons cannot always be replicated. *Phys. Lett. A* **92**(7), 321 (1982)
4. B. Kanseri, T. Iskhakov, I. Agafonov, M. Chekhova, G. Leuchs, Three-dimensional quantum polarization tomography of macroscopic Bell states. *Phys. Rev. A* **85**(2), 022126 (2012)

5. M. Cramer, M.B. Plenio, S.T. Flammia, R. Somma, D. Gross, S.D. Bartlett, O. Landon-Cardinal, D. Poulin, Y.K. Liu, Efficient quantum state tomography. *Nat. Commun.* **1**(9), 149 (2010)
6. M. Hofheinz, H. Wang, M. Ansmann, R.C. Bialczak, E. Lucero, M. Neeley, A.D. O'Connell, D. Sank, J. Wenner, J.M. Martinis, A.N. Cleland, Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**(7246), 546 (2009)
7. K. Resch, P. Walther, A. Zeilinger, Full characterization of a three-photon Greenberger-Horne-Zeilinger state using quantum state tomography. *Phys. Rev. Lett.* **94**(7), 070402 (2005)
8. M. Beck, C. Dorner, I. Walmsley, Joint quantum measurement using unbalanced array detection. *Phys. Rev. Lett.* **87**(25), 253601 (2001)
9. D.F.V. James, P.G. Kwiat, W.J. Munro, A.G. White, Measurement of qubits. *Phys. Rev. A* **64**(5), 052312 (2001)
10. M.G. Raymer, M. Beck, D. McAlister, Complex wave-field reconstruction using phase-space tomography. *Phys. Rev. Lett.* **72**(8), 1137 (1994)
11. D. Smithey, M. Beck, M. Raymer, A. Faridani, Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**(9), 1244 (1993)
12. K. Vogel, H. Risken, Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **40**(5), 2847 (1989)
13. M. Agnew, J. Leach, M. McLaren, F.S. Roux, R.W. Boyd, Tomography of the quantum state of photons entangled in high dimensions. *Phys. Rev. A* **84**(6), 062101 (2011)
14. J.S. Lundeen, B. Sutherland, A. Patel, C. Stewart, C. Bamber, Direct measurement of the quantum wavefunction. *Nature* **474**(7350), 188 (2011)
15. J.Z. Salvail, M. Agnew, A.S. Johnson, E. Bolduc, J. Leach, R.W. Boyd, Full characterization of polarization states of light via direct measurement. *Nature Photon.* **7**(4), 316 (2013)
16. M. Malik, M. Mirhosseini, M.P.J. Lavery, J. Leach, M.J. Padgett, R.W. Boyd, Direct measurement of a 27-dimensional orbital-angular-momentum state vector. *Nat. Commun.* **5**, 3115 (2014)
17. J. Fischbach, M. Freyberger, Quantum optical reconstruction scheme using weak values. *Phys. Rev. A* **86**(5), 052110 (2012)
18. S. Wu, State tomography via weak measurements. *Sci. Rep.* **3**, (2013)
19. A. Di Lorenzo, Sequential measurement of conjugate variables as an alternative quantum state tomography. *Phys. Rev. Lett.* **110**(1), 010404 (2013)
20. Y. Aharonov, D. Albert, L. Vaidman, How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.* **60**(14), 1351 (1988)
21. J.S. Lundeen, K. Resch, Practical measurement of joint weak values and their connection to the annihilation operator. *Phys. Lett. A* **334**, 337 (2005)
22. J. Dressel, S. Agarwal, A.N. Jordan, Contextual values of observables in quantum measurements. *Phys. Rev. Lett.* **104**(24), 240401 (2010)
23. P. Dixon, D. Starling, A. Jordan, J. Howell, Ultrasensitive beam deflection measurement via interferometric weak value amplification. *Phys. Rev. Lett.* **102**(17) (2009)
24. N. Ritchie, J. Story, R. Hulet, Realization of a measurement of a “weak value”. *Phys. Rev. Lett.* **66**(9), 1107 (1991)
25. O. Hosten, P. Kwiat, Observation of the spin hall effect of light via weak measurements. *Science* **319**(5864), 787 (2008)
26. O.S. Magaña-Loaiza, M. Mirhosseini, B. Rodenburg, R.W. Boyd, Amplification of angular rotations using weak measurements. *Phys. Rev. Lett.* **112**(20), 200401 (2014)
27. N. Brunner, C. Simon, Measuring small longitudinal phase shifts: weak measurements or standard interferometry? *Phys. Rev. Lett.* **105**(1), 010405 (2010)
28. N. Brunner, V. Scarani, M. Wegmüller, M. Legré, N. Gisin, Direct measurement of superluminal group velocity and signal velocity in an optical fiber. *Phys. Rev. Lett.* **93**(20), 203902 (2004)
29. A. Feizpour, X. Xing, A.M. Steinberg, Amplifying single-photon nonlinearity using weak measurements. *Phys. Rev. Lett.* **107**(13), 133603 (2011)

30. M. Mirhosseini, O.S. Magaña-Loaiza, S.M. Hashemi Rafsanjani, R.W. Boyd, Compressive direct measurement of the quantum wave function. *Phys. Rev. Lett.* **113**(9), 090402 (2014)
31. C. Bamber, B. Sutherland, A. Patel, C. Stewart, J.S. Lundeen, Measurement of the transverse electric field profile of light by a self-referencing method with direct phase determination. *Opt. Express* **20**(3), 2034 (2012)
32. J. Lundeen, C. Bamber, Procedure for direct measurement of general quantum states using weak measurement. *Phys. Rev. Lett.* **108**(7) (2012)
33. C. Bamber, J.S. Lundeen, Observing dirac's classical phase space analog to the quantum state. *Phys. Rev. Lett.* **112**, 070405 (2014)
34. G.S. Thekkadath, L. Giner, Y. Chalich, M.J. Horton, J. Bunker, J.S. Lundeen, Direct measurement of the density matrix of a quantum system. *Phys. Rev. Lett.* **117**(12), 120401 (2016)
35. G.A. Howland, D.J. Lum, J.C. Howell, Compressive wavefront sensing with weak values. *Opt. Express* **22**(16), 18870 (2014)
36. E. Yao, S. Franke-Arnold, J. Courtial, S. Barnett, M. Padgett, Fourier relationship between angular position and optical orbital angular momentum. *Opt. Express* **14**(20), 9071 (2006)
37. B. Jack, M.J. Padgett, S. Franke-Arnold, Angular diffraction. *New J. Phys.* **10**(10), 103013 (2008)
38. M. Mirhosseini, O.S. Magaña-Loaiza, C. Chen, S.M.H. Rafsanjani, R.W. Boyd, Wigner distribution of twisted photons. *Phys. Rev. Lett.* **116**(13), 130402 (2016)
39. M. Mirhosseini, M. Malik, Z. Shi, R.W. Boyd, Efficient separation of the orbital angular momentum eigenstates of light. *Nat. Commun.* **4**, 2781 (2013)
40. M. Mirhosseini, O.S. Magaña-Loaiza, M.N. O'Sullivan, B. Rodenburg, M. Malik, M.P.J. Lavery, M.J. Padgett, D.J. Gauthier, R.W. Boyd, High-dimensional quantum cryptography with twisted light. *New J. Phys.* **17**(3), 033033 (2015)
41. A. Shabani, R.L. Kosut, M. Mohseni, H. Rabitz, M.A. Broome, M.P. Almeida, A. Fedrizzi, A.G. White, Efficient measurement of quantum dynamics via compressive sensing. *Phys. Rev. Lett.* **106**(10), 100401 (2011)
42. G. Howland, J. Howell, Efficient high-dimensional entanglement imaging with a compressive-sensing double-pixel camera. *Phys. Rev. X* **3**(1), 011013 (2013)
43. W.T. Liu, T. Zhang, J.Y. Liu, P.X. Chen, J.M. Yuan, Experimental quantum state tomography via compressed sampling. *Phys. Rev. Lett.* **108**(17), 170403 (2012)
44. D. Gross, Y.K. Liu, S.T. Flammia, S. Becker, J. Eisert, Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105**(15), 150401 (2010)
45. O. Katz, Y. Bromberg, Y. Silberberg, Compressive ghost imaging. *Appl. Phys. Lett.* **95**(13), 131110 (2009)
46. R.G. Baraniuk, Single-pixel imaging via compressive sampling. *IEEE Sig. Process Mag.* (2008)
47. J. Romberg, Imaging via compressive sampling. *IEEE Sig. Process Mag.* **25**(2), 14 (2008)
48. E. Candes, J. Romberg, Sparsity and incoherence in compressive sampling. *Inverse Probl.* **23**(3), 969 (2007)
49. P. Zerom, K.W.C. Chan, J.C. Howell, R.W. Boyd, Entangled-photon compressive ghost imaging. *Phys. Rev. A* **84**(6), 061804 (2011)
50. K. Lyons, J. Dressel, A.N. Jordan, J.C. Howell, P.G. Kwiat, Power-recycled weak-value-based metrology. *Phys. Rev. Lett.* **114**(17), 170801 (2015)
51. C. Bamber, J.S. Lundeen, Observing Dirac's classical phase space analog to the quantum state. *Phys. Rev. Lett.* **112**(7), 070405 (2014)
52. Z. Shi, M. Mirhosseini, J. Margiewicz, M. Malik, F. Rivera, Z. Zhu, R.W. Boyd, Scan-free direct measurement of an extremely high-dimensional photonic state. *Optica* **2**(4), 388–392 (2015)
53. D.F. McAlister, M. Beck, L. Clarke, A. Mayer, M.G. Raymer, Optical phase retrieval by phase-space tomography and fractional-order fourier transforms. *Opt. Lett.* **20**(10), 1181 (1995)
54. T. Durt, B.G. Englert, I. Bengtsson, On mutually unbiased bases. *Int. J. Quantum Inf.* **08**(04), 535 (2010)

55. H.F. Hofmann, Complex joint probabilities as expressions of reversible transformations in quantum mechanics. *New J. Phys.* **14**(4), 043031 (2012)
56. E. Arthurs, J. Kelly, On the simultaneous measurement of a pair of conjugate observables. *Bell System Tech. J.* **44**, 725 (1965)
57. J.H. Shapiro, S.S. Wagner, Phase and amplitude uncertainties in heterodyne detection. *IEEE J. Quantum Electron.* **20**(7), 803 (1984)
58. U. Leonhardt, H. Paul, Phase measurement and q function. *Phys. Rev. A* **47**(4), R2460 (1993)
59. U. Leonhardt, Quantum-state tomography and discrete Wigner function. *Phys. Rev. Lett.* **74**(21), 4101 (1995)
60. U. Leonhardt, Discrete Wigner function and quantum-state tomography. *Phys. Rev. A* **53**(5), 2998 (1996)
61. E. Haapasalo, P. Lahti, J. Schultz, Weak versus approximate values in quantum state determination. *Phys. Rev. A* **84**(5), 052107 (2011)
62. G. Vallone, D. Dequal, Strong measurements give a better direct measurement of the quantum wave function. *Phys. Rev. Lett.* **116**(4), 040502 (2016)
63. R. Okamoto, M. Iefuji, S. Oyama, K. Yamagata, H. Imai, A. Fujiwara, S. Takeuchi, Experimental demonstration of adaptive quantum state estimation. *Phys. Rev. Lett.* **109**, 130404 (2012)
64. D.H. Mahler, L.A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, A.M. Steinberg, Adaptive quantum state tomography improves accuracy quadratically. *Phys. Rev. Lett.* **111**, 183601 (2013)
65. E. Bolduc, G. Gariepy, J. Leach, Direct measurement of large-scale quantum states via expectation values of non-Hermitian matrices. *Nat. commun.* **7**, (2016)
66. U. Leonhardt, *Measuring the Quantum State of Light. Cambridge Studies in Modern Optics* (Cambridge University Press, 1997)

Chapter 3

Nonlinear Interactions and Non-classical Light



Dmitry V. Strekalov and Gerd Leuchs

Abstract The term *non-classical* concerns light whose properties cannot be explained by classical electrodynamics and which requires invoking quantum principles to be understood. Its existence is a direct consequence of field quantization; its study is a source of our understanding of many quantum phenomena. Non-classical light also has properties that may be of technological significance. We start this chapter by discussing the definition of non-classical light and basic examples. Then some of the most prominent applications of non-classical light are reviewed. After that, as the principal part of our discourse, we review the most common sources of non-classical light. We will find them surprisingly diverse, including physical systems of various sizes and complexity, ranging from single atoms to optical crystals and to semiconductor lasers. Putting all these dissimilar optical devices in the common perspective we attempt to establish a trend in the field and to foresee the new cross-disciplinary approaches and techniques of generating non-classical light.

3.1 Introduction

3.1.1 Classical and Non-classical Light

In historical perspective, light doubtlessly is among the most classical phenomena of physics. The oldest known treatise on this subject, “Optics” by Euclid, dates back to approximately 300 B.C. Yet in contemporary physics, light is one of the strongest manifestations of quantum. Electromagnetic field quanta, the *photons*, are certainly real: they can be emitted and detected one by one, delivering discrete portions of energy and momentum, and in this sense may be viewed as particles of light. At

D. V. Strekalov (✉) · G. Leuchs

Max Planck Institute for the Science of Light, Staudstraße 2, 90158 Erlangen, Germany
e-mail: strelkov@yahoo.com

G. Leuchs
e-mail: gerd.leuchs@mpl.mpg.de

the same time, wave properties of light are most readily observed in diffraction and interference experiments.

Before the quantum mechanical principle of *duality* was understood, this twofold nature of light has lead to curious oscillations in its understanding by scientists and philosophers starting from antiquity. Pythagoras believed light to consist of moving particles, but Aristotle later compared it to ocean waves. Much later Sir Isaac Newton has revived the concept of corpuscles, which again yielded the ground to the wave theory when interference and diffraction were discovered. Then the sum of evidence for each, the wave and the corpuscular nature of light, became undeniable and quantum optics emerged.

So what is classical and what is non-classical light? Before going on we note that nature is as it is, and any distinction between quantum and classical is somewhat artificial and is merely a result of our desire to describe nature by models with which we can make quantitative predictions. There is a whole class of models labeled classical because they have or could have been formulated before the invention of quantum physics. Nevertheless it can help our perception to see how far a particular model, be it in the classical or in the quantum class, can be stretched before it fails making correct predictions. No matter which type of light one studies, it seems obvious that one can in principle find some departure between the model and the observation.

It appears logical to call “classical” the phenomena that can be quantitatively described without invoking quantum mechanics, e.g. in terms of Maxwell’s equations. Interference and diffraction are obvious examples from classical optics. Somewhat less obvious examples are “photon bunching”, Hanbury Brown and Twiss type interference of thermal light, and a few other phenomena that occasionally raise the quantum-or-classical debate in conference halls and in the literature.

Conversely, non-classical (quantum) are those phenomena that can *only* be described in quantum mechanics. It should be noted that in many cases it is *convenient* to describe classical light in terms of quantum optics, which, however, does not make it non-classical in the sense mentioned above. This is done in order to use the same language for classical and quantum phenomena analysis. As stated above, nature does not make this distinction. It is our choice if we employ as much as possible a classical model with a limited applicability.

One of the most useful quantum versus classical distinction criteria is based on the various correlation functions of optical fields. Such correlation functions are computed by averaging the observables over their joint probability distribution. For a simple example let us consider a normalized auto-correlation function of light intensity $I(t) = E^*(t)E(t)$, which is known as the Glauber correlation function*

$$g^{(2)}(\tau) = \frac{\langle E^*(t)E^*(t + \tau)E(t)E(t + \tau) \rangle}{\langle E^*(t)E(t) \rangle^2}. \quad (3.1)$$

Using the Cauchy-Schwarz inequality, it is easy to see [1, 2] that $g^{(2)}(0) \geq 1$. Smaller values for this observable are impossible in classical optics, but they do occur in

*We salute to Roy Jay Glauber who pioneered quantum optics and the quantum version of field correlation functions. He passed away on 26 December 2018 while we were reading the proofs.

nature, e.g. for photon number states and for amplitude-squeezed light. Therefore *antibunching* [3–7] $g^{(2)}(0) < 1$ can be taken as a sufficient but not necessary criterion for non-classical light.¹ A similar argument can be made for the intensity correlation (3.1) as a function of spatial coordinates instead of time. In this case inequalities similar to Cauchy-Schwarz lead to such non-classicality criteria as Bell’s inequalities violation [9–12] and negative conditional Von Neumann entropy [13, 14].

Another criterion, likewise sufficient but not necessary, is the negativity of the phase space distribution function [15]. In quantum mechanics such distribution functions can be introduced with limiting cases being the Wigner function, or the Glauber-Sudarshan P-function. Negative, complex or irregular values of these functions can also be used as indications of non-classical light [16–19]. We note that one of the limiting cases, the Q-function (also referred to as the Husimi function), is always regular.

Both criteria are sufficient but not necessary as can be seen by the following examples: Photon number states are non-classical according to both criteria; amplitude squeezed states are non-classical according to the first but not the second criterion; superpositions of coherent states, so-called cat states [20, 21],² are non-classical according to the second but not the first criterion.

The qualifier “not necessary” in the above criteria is essential. Currently we know of no simple general criterion, which is sufficient *and* necessary. But we can make the following statement: classical states can involve either no fluctuations at all, or only statistical fluctuations. In quantum physics such states either do not exist, or they are described as mixed states. Pure quantum states exhibit a so-called quantum uncertainty, which results in the projection noise when measured. A classical stochastic model can describe some aspects of a quantum uncertainty. But such models are always limited. Allowing for all possible experimental scenarios, a pure quantum state can never be described by one and the same classical stochastic model. In this sense coherent states, being pure quantum states in their own right, have to be classified as non-classical. This statement calls for a more detailed justification, which is provided in the Appendix in the end of this chapter.

3.1.2 Types of Non-classical Light

A key concept in the following discussion will be an optical mode. This concept is fundamental to electromagnetic field quantization: the spatio-temporal character of a mode is described by a real-valued mode function, which is a function of position and time and is normalized. In classical physics this amplitude function is multiplied by

¹In practice, it is often convenient to measure autocorrelation function (3.1) using a beam splitter and a pair of detectors. The same or similar set up can be used for measuring a *cross-correlation* function of two optical modes. Note that this measurement yields a different observable whose classical range is $g_{12}^{(2)}(0) > 0.5$ [8].

²The cat states, named after Schrödinger’s cat, are superpositions of two out-of-phase macroscopic ($|\alpha| \gg 1$) coherent states, e.g. $|\Psi\rangle_{\text{cat}} \propto |\alpha\rangle + |-\alpha\rangle$.

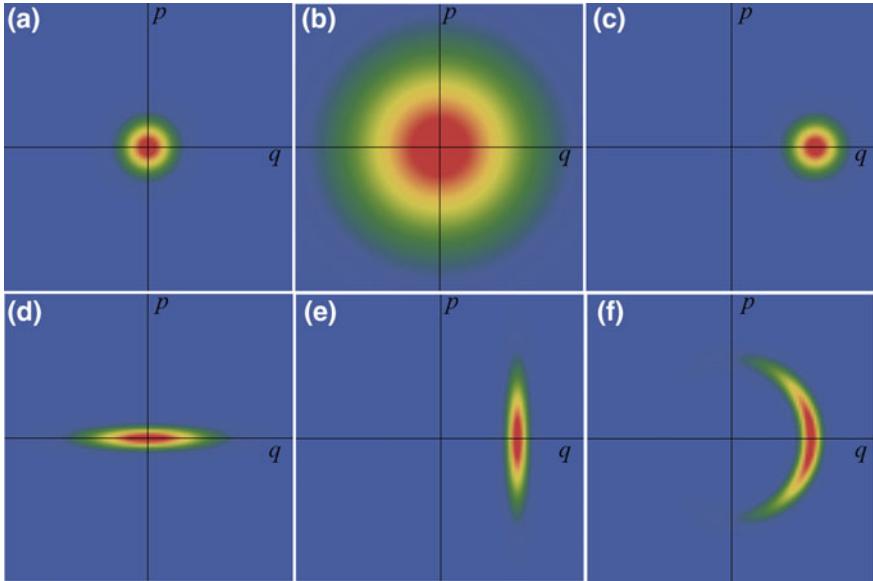


Fig. 3.1 Phase space diagrams for symmetric ordering (Wigner functions) for some classical and non-classical states: vacuum (a), thermal (b), coherent (c), squeezed vacuum (d), quadrature squeezed light (e), and photon-number squeezed light (f). Mean photon number for states (b), (c), (e) and (f) is $\langle \hat{N} \rangle = 6$, for (d) $\langle \hat{N} \rangle = 1$. Squeezing parameter for states (d)–(f) is 7.66 dB

a complex number describing amplitude and phase, or alternatively two orthogonal field quadratures. In quantum optics these amplitudes are described by operators; superpositions of the photon creation and annihilation operators, to be precise.

Quantum-statistical properties of single-mode light are conveniently illustrated by phase space diagrams where the optical state Wigner function $W(p, q)$ is plotted against the canonical harmonic oscillator coordinates p and q . We recall [1] that for an optical mode with central frequency ω the corresponding quadrature operators are related to the photon creation and annihilation operators a^\dagger and a as

$$\hat{p} = -i\sqrt{\hbar\omega/2}(a - a^\dagger), \quad \hat{q} = -i\sqrt{\hbar/(2\omega)}(a + a^\dagger). \quad (3.2)$$

In Fig. 3.1 we show some examples of classical and non-classical light phase diagrams. Here (a) is the vacuum state, and (b), (c) are the thermal and coherent states with the mean photon number $\langle \hat{N} \rangle = \langle a^\dagger a \rangle = 6$, respectively (the plots are scaled for $\hbar\omega = 1$). The diagram (d) represents squeezed vacuum with $\langle \hat{N} \rangle = 1$. For the squeezed vacuum states the mean photon number is uniquely related to squeezing; our example $\langle \hat{N} \rangle = 1$ requires $20 \lg(e) \operatorname{arcsinh}(\langle \hat{N} \rangle^{1/2}) \approx 7.66$ dB of squeezing. The diagrams (e) and (f) show the quadrature and amplitude (or photon-number) squeezed states, respectively, with the same squeezing factor as in (d) and the same mean photon number as in (b) and (c).

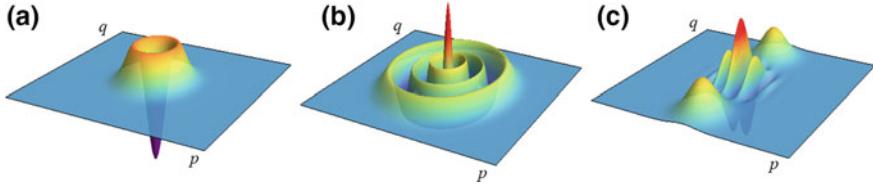


Fig. 3.2 Phase space diagrams for non-Gaussian states: single-photon (a) and six-photon (b) Fock states, and a cat state corresponding to superposition of two $\langle N \rangle = 6$ coherent states such as shown in Fig. 3.1c

The thermal state represented in Fig. 3.1b is clearly classical. The vacuum state (a) and the coherent state (c), which is also called displaced vacuum state, are said to be at the quantum-classical boundary. They do not violate either of the two criteria formulated above, but the boundary value $g^{(2)}(0) = 1$ implies that the optical field does not fluctuate, see (3.1). Then the Poissonian statistics of photocounts, observed with a coherent field, must be attributed to stochastic character of the detection process. But if this were the case, it would not be possible to observe sub shot noise correlation of two independent detectors' signals (such as measured in two-mode squeezing experiments) even with 100%-efficient photodetectors. Therefore one is led to a conclusion that detection of coherent light cannot be fully described in the semiclassical approximation, and in this sense such light is non-classical. Despite these notational difficulties coherent states as any other pure quantum state can be used as resource for optical quantum engineering, such as in quantum key distribution.

Wigner functions for the classical and non-classical states shown in Fig. 3.1 are symmetric or distorted Gaussian. Other non-classical states may be non-Gaussian, and clearly displaying Wigner function negativity. Such are the Fock states and the already mentioned cat states, shown in Fig. 3.2. Their Wigner functions are respectively

$$W_{\text{Fock}} = \frac{2}{\pi} (-1)^N L_N(2p^2 + 2q^2) e^{-p^2 - q^2}, \quad (3.3)$$

$$W_{\text{cat}} = \frac{e^{-p^2 - (q - |\alpha|)^2} + e^{-p^2 - (q + |\alpha|)^2} + 2 \cos(2p|\alpha|) e^{-p^2 - q^2}}{\pi \sqrt{2(1 + e^{-2|\alpha|^2})}},$$

where L_N is the N -th order Laguerre polynomial. It is interesting to observe that the Wigner function of a photon-number eigenstate indicates a non-zero quasi-probability for the mode to be found in a *different* number-state, and in fact reaches the maximum for the vacuum state $|N\rangle = 0$. This reminds us to be cautious with physical interpretations of quasi-probability functions such as the Wigner function.

The standard way of measuring and gauging the intensity fluctuations is to split the optical beam with a 50/50 beam splitter and either subtract or multiply the photocurrents of two detectors placed at each output. Time dependence may be obtained

by introducing a variable optical or electronic delay in one channel. This approach would identify the photon-number squeezed state shown in Fig. 3.1f as non-classical, but would not reveal the non-classical properties of the squeezed states in diagrams (d) and (e). In fact, an ensemble of measurements on such identically prepared states (which is often equivalent to a time-sequential measurement on one system) would show excessive photon-number fluctuation, above the shot noise limit. To measure the quadrature squeezing one has to set up a measurement sensitive to the Wigner function projection on the squeezed quadrature, rather than in the radial direction. This can be achieved in a heterodyne measurement, when a coherent local oscillator field is injected into the unused port of the beam splitter. Changing the local oscillator phase one can chose the projection direction. The right choice of the phase leads to a sub shot noise measurement revealing the non-classicality: $g^{(2)}(0) < 1$. The same situation can be described in a different language, by saying that the beam splitter transforms the input (e.g. quadrature-squeezed and coherent) modes to output modes, both of which are photon-number squeezed, or anti-bunched.

Non-classical phenomena in two or more optical modes are usually associated with the term *entanglement*. A quantum optical system comprising modes labeled A and B is said to be entangled if its wave function does not factorize: $|\Psi\rangle_{AB} \neq |\Psi\rangle_A \otimes |\Psi\rangle_B$. This concept can be applied to systems of more than two modes, in which case one has multipartite entanglement. Entangled states can also be described in density operators notation, which allows to consider the states that are not quantum-mechanically pure.

Perhaps the most common examples of entangled states in optics are so-called Bell states of a polarization-entangled photon pair³

$$\begin{aligned} |\Psi^{(-)}\rangle_{AB} &= (|\uparrow\rangle_A |\leftrightarrow\rangle_B - |\leftrightarrow\rangle_A |\uparrow\rangle_B) / \sqrt{2}, \\ |\Psi^{(+)}\rangle_{AB} &= (|\uparrow\rangle_A |\leftrightarrow\rangle_B + |\leftrightarrow\rangle_A |\uparrow\rangle_B) / \sqrt{2}, \\ |\Phi^{(\pm)}\rangle_{AB} &= (|\uparrow\rangle_A |\downarrow\rangle_B \pm |\leftrightarrow\rangle_A |\leftrightarrow\rangle_B) / \sqrt{2}, \end{aligned} \quad (3.4)$$

and frequency entangled states such as

$$|\Psi\rangle_{AB} = \int |\omega_A + v\rangle_A |\omega_B - v\rangle_B F(v) dv, \quad (3.5)$$

where A and B designate spatial modes. It is also possible to have a photon pair simultaneously entangled in *both* polarization and frequency (or equivalently, time) [22].

Not only a single pair of photons may be entangled. It is also possible to create an entangled state with larger certain or uncertain photon numbers. One of the examples is vacuum entangled with a Fock state $|\Psi\rangle_{AB} = (|N\rangle_A |0\rangle_B + |0\rangle_A |N\rangle_B) / \sqrt{2}$,

³We use a notation where a vertical or horizontal arrow represents one of the two orthogonal linear polarizations, and subscripts A and B one of the two spatial modes. Hence we work in four-dimensional Hilbert space where single-photon base states can be mapped as follows: $|\uparrow\rangle_A \rightarrow |1, 0, 0, 0\rangle$, $|\leftrightarrow\rangle_A \rightarrow |0, 1, 0, 0\rangle$, $|\downarrow\rangle_B \rightarrow |0, 0, 1, 0\rangle$, $|\leftrightarrow\rangle_B \rightarrow |0, 0, 0, 1\rangle$.

dubbed “NOON-state” [22–24]. Macroscopical states can be entangled not only in photon numbers, but also in the canonical coordinates (quadratures) p and q . To distinguish it from the entanglement in the discrete photon numbers, quadrature entanglement is also called continuous-variable entanglement. It can be generated e.g. by combining two squeezed vacuum states on a beam splitter [25, 26], and forms a foundation for continuous-variable quantum information processing [27] and quantum state teleportation [28]. The discrete and continuous variable descriptions correspond to expanding the wave function in two different bases. One and the same state can be represented by either one of them. On the practical side: photon number resolving detectors measure in terms of the discrete Fock state basis and homodyne detection measures in terms of quadrature basis.

Graphic representation of two- or multi-mode non-classical states on a phase diagram is more complicated than for a single mode. In general, it requires as many diagrams as there are modes, with a color coding indicating the quantum-correlated sub-spaces *within* each diagram [29]. This appears to allow for a better photon localization in phase space than is permitted by Heisenberg uncertainty. However the uncertainty principle is not really violated, because the localization occurs in superpositions of quadratures of different field operators that do commute, e.g. $[x_1 - x_2, p_1 + p_2] = 0$, giving rise to Einstein-Podolski-Rosen correlations.

The quantum state of any one mode of a system comprising many modes and described by a multipartite state can be found by taking a trace over the unobserved modes. If initially the entire system was in a pure entangled state, the single mode sub-system will be found in a mixed state, as can e.g. be seen starting from (3.4) and (3.5). This also can be understood following a von Neumann entropy analysis [13]. Indeed, if a bipartite system is in a pure state with $S = 0$, and the conditional entropy is negative $S_{B|A} < 0$ because of entanglement, then the entropy of a sub-system is positive, $S_A = S - S_{B|A} > 0$, which means that it is in a mixed state. Remarkably, in some cases this does not preclude this mode from being in a non-classical state. For example, the twin beams of an optical parametric oscillator (OPO) that are well-known to be quantum-correlated (or two-mode squeezed), are predicted [30–32] and demonstrated [33] to be also single-mode squeezed when the OPO is well above the threshold. In this case one finds a mixed squeezed state which occupies a larger area in phase space than required by the uncertainty relation.

It should be noted that two-mode quantum correlation and single spatial mode non-classical photon statistics are often viewed as two sides of the same coin. This affinity, emphasized by the use of the term “two-mode squeezing” in analogy with the “two-mode entanglement”, arises from the simplicity of conversion between these types of photon statistics. The conversion is performed with a linear beamsplitter, and can be elegantly described [34, 35] by an SU(2) operator converting two input states to two output states. This operation leads to a conversion of phase fluctuations into amplitude fluctuations, and of two-mode entanglement into single-mode squeezing [24, 29, 36].

A special case of two-mode squeezing is realized when the modes are associated with orthogonal polarizations of the same optical beam. Just like a spatial mode can be associated with any function from an orthogonal set of Helmholtz equation

solutions (e.g. Laguerre-Gauss modes), here we are free to chose any polarization basis to designate polarization modes. It is often convenient to chose a linear basis (x, y) . In this case polarization Stokes operators are introduced as

$$\begin{aligned}\hat{S}_1 &= a_x^\dagger a_x - a_y^\dagger a_y, \\ \hat{S}_2 &= a_x^\dagger a_y + a_y^\dagger a_x, \\ \hat{S}_3 &= i(a_y^\dagger a_x - a_x^\dagger a_y).\end{aligned}\quad (3.6)$$

Like the canonical coordinate or quadrature operators (3.2), Stokes operators do not commute. They too span a phase space (three-dimensional instead of two-dimensional, since we now have two independent polarization modes) where a pure state occupies the minimum volume allowed by the uncertainty relations. Its shape, however, can be distorted—squeezed. For example, squeezing in the S_1 quadrature can be observed as sub shot noise fluctuations of the difference of the currents generated by two photo detectors set to measure optical powers in the x and y linear polarizations.

With increasing the number of modes, which in optics may be associated with the Hilbert space dimension, the list of possible non-classical states rapidly grows. Some examples are the entangled states of multiple photons in different modes, such as optical Greenberger-Horne-Zeilinger (GHZ) states [37, 38], W states [39–41], as well as cluster [42] and graph [43] states, Smolin states [44] and others.

In quantum communications, higher-dimensional entanglement provides a higher information capacity [45–48]. From a fundamental point of view, higher-dimensional entanglement leads to stronger violations of generalized Bell's inequalities [49]. This has been experimentally demonstrated in a 16-dimentional Hilbert space spanned by the optical polarization states [50] and in a 12-dimensional Hilbert space spanned by the optical orbital angular momentum (OAM) states [51].

Entanglement in the Hilbert space spanned by OAM states [52] is a relatively novel and very promising approach to generating multi-mode entanglement. Two-photon entanglement in 100×100 —dimensional space was demonstrated following this approach [53], as was the *four-photon* entanglement [54].

Entanglement of a 100, and with certain allowances of even a 1000 optical modes based on polarization, rather than spacial, variables has also been theoretically discussed [55] and shown to be within reach with the existing technology. However applying the entanglement metrics [56] such as negativity [57] or concurrence [58, 59] shows that such states are very close to classical light.

Let us now review some of the practical applications that make non-classical light such an important topic in optics.

3.1.3 Applications of Non-classical Light

The fact that light can possess non-classical properties that can only be explained in the framework of quantum mechanics is remarkable and important for our understanding of Nature. Besides that, nonclassical light can have useful technological applications.

Absolute Calibration of Light Detectors and Sources

Perhaps the oldest application of non-classical light, proposed back in 1969–1970 [60, 61] and further developed by Klyshko in 1980 [62], is the absolute calibration of the quantum efficiency of photon counting detectors. The concept underlying this method is very simple. Suppose a process generating photon pairs, such as spontaneous parametric down conversion (SPDC), produces N signal and idler photon pairs per second. The photons are sent into photon counting detectors with quantum efficiencies η_1 for the signal channel and η_2 for the idler channel. Imperfect detection $\eta_{1,2} < 1$ leads to *random* loss of photons in both detectors. Then the mean values for the number of photocounts $N_{1,2}$ and for coincidence counts N_c are found as $N_{1,2} = N\eta_{1,2}$ and $N_c = N\eta_1\eta_2$. Therefore both quantum efficiencies can be inferred by counting the individual and coincidence detections: $\eta_{1,2} = N_{2,1}/N_c$. In practical applications one needs to account for multiple pairs occasionally generated in SPDC during a coincidence window, dark noise and dead time of the detectors, and other factors that make the calibration formula and procedure more complicated [63, 64].

A single-detector implementation of this technique was also discussed in [62]. This requires a photon number resolving detector collecting all of SPDC light (both the signal and the idler components) near degeneracy. This technique is based on comparing the single- and double-photon detection probabilities. Like the two-detector method, it also received further development [65, 66].

Another possibility that was mentioned in [62] is calibration of photo detectors operating in the photo current (continuous) regime instead of photon counting (Geiger) regime. In this case, a correlation function of two photo currents is used instead of the coincidence counting rate [67]. Note that since the discrete character of photo detections is no longer required, this method allows for using the two-mode squeezed light instead of two-photon SPDC light [67, 68]. A multimode version of this method was used for calibration of CCD cameras [69].

Similarly to spontaneous emission by excited atoms, SPDC can be viewed as amplification of vacuum uncertainty of the optical field. This vacuum uncertainty is often referred to as *vacuum fluctuations*. But strictly speaking this is a time independent uncertainty which is stochastically projected on a single value when measured. When repeating the process of state preparation and measurement many times, the uncertainty is transformed into an apparent fluctuation. Note that a measurement does not necessarily involve the action of a human experimenter. Coupling the system under study to some environment which then loses coherence (i.e. which decoheres) has the same effect.

The vacuum uncertainty has a spectral brightness⁴ of $S_{\text{vac}} = \hbar c \lambda^{-3}$ [70]. Since parametric amplification of weak signals is linear, it is possible to perform absolute calibration of a light source directly in the units of S_{vac} by seeding its light into a parametric amplifier and comparing the emitted parametric signals with and without seeding [71].

Sub Shot Noise Measurements

We already noted that the power fluctuations, or noise of non-classical light may be reduced below the classical shot noise limit. This effect may be used for low-noise measurements of a variable of interest. First application of squeezed vacuum for sub shot noise interferometric phase measurements has been demonstrated [72] already in 1987, followed by another publication from a different group [73]. In these works squeezed vacuum was generated in a degenerate-wavelength OPO pumped below the threshold by the second harmonic of the coherent laser light used in the interferometer. This technique, now commonly used in the field, fixes the frequency and phase relation between the coherent signal and squeezed vacuum. Injecting the squeezed vacuum into a dark port of an interferometer reduces the signal fluctuations below the shot noise by an amount which depends on the degree of squeezing. A reduction figure of 3.5 dB was reached with this approach in the GEO 600 setup of the LIGO project [74], see Fig. 3.3. In this case the state-of-the-art 10 dB squeezed vacuum resource was used. However, imperfect transmission of the complex multi-path interferometer $\eta = 0.62 < 1$ increased the observed signal variance (i.e., noise) from the squeezed vacuum source value $V_{\text{sqz}}^{(0)} = 0.1$ to $V_{\text{sqz}} = \eta \times V_{\text{sqz}}^{(0)} + (1 - \eta) \times 1 = 0.44$. This calculated variance agrees well with the reported 3.5 dB of shot noise suppression.

Besides interferometry, non-classical light can facilitate sub shot noise measurements in spectroscopy [75, 76] and in biological research [77]. On the other hand, strong intensity fluctuations can enhance the two-photon absorption in atoms and other systems, compared with light of the same average intensity but Poisson or sub-Poisson fluctuations. Theoretical analysis of this phenomenon in two-photon and squeezed light predicts a linear (rather than quadratic) dependence of the absorption rate on the optical intensity for weak fields, the possibility of a decreasing absorption rate with increasing intensity, and a significant differences between absorption rates for the phase- and amplitude-squeezed beams of the same intensity [78, 79]. Further theoretical analysis including the second harmonic generation is provided in [80].

Two-photon absorption of non-classical light has been observed with cesium [81] and rubidium [82] atoms. In both cases atomic two-photon transitions were excited by non-degenerate squeezed light generated in an OPO cavity. Excitation rate scaling as the power 1.3 (instead of 2) of the light intensity was observed in [81]. Conversely, it is possible to characterize photon bunching by observing two-photon response in semiconductors [83].

⁴We recall that spectral brightness, determining the mean number of photons per mode, in free space is measured in terms of light intensity emitted into a unity solid angle per unity frequency bandwidth.

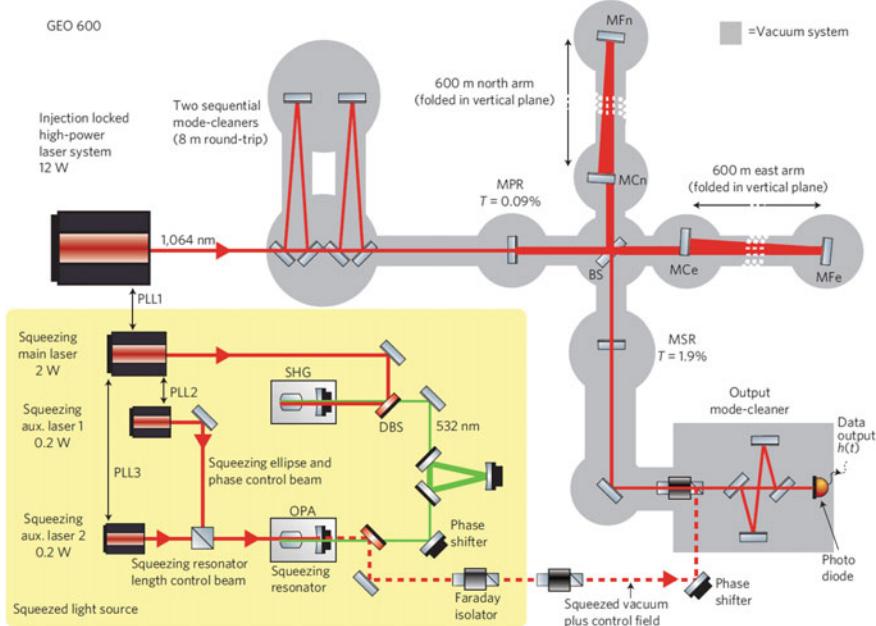


Fig. 3.3 A complex LIGO interferometer uses a squeezed vacuum input to reduce the measurement noise below the shot-noise limit. Reprinted from [74]

Speaking of spectroscopy, we must mention yet another application of non-classical light, not related to noise reduction but remarkable nonetheless. In this application strongly non-degenerate SPDC light propagates in a nonlinear interferometer filled with a sample of refractive material [84]. As expected, a strong dispersion in e.g. infrared range is indicated by the characteristic distortion patterns of interference fringes in the infrared (idler) port. However it also leads to similar distortions arising in the signal port, which allows for performing infrared spectroscopy using visible light optics and detectors.

High-Resolution Imaging

The term “imaging” may refer to both creating and reading of patterns, as well as to optical detection of small displacements. All these functionalities have been shown to benefit from applications of non-classical light. Creating lithographic images with higher than diffraction-limited resolution has been proposed in year 2000 [85]. This proposal is based on using photo-polymers sensitive to N -photon absorption in conjunction with already mentioned entangled NOON states. It was theoretically shown that using these states in a Mach-Zehnder interferometer one can generate $\lambda/(2N)$ -spaced fringes of the N th order intensity distribution $\langle I^N \rangle$ that would imprint in the polymer. It should be noted that even with classical light the N -photon material response by itself provides a \sqrt{N} reduction of the optical point-spread function.

With special modulation techniques this reduction factor can be further pushed to reach the quantum limit of N [86]. Therefore the practical benefit of the quantum lithography proposal turned out to be limited. However its originality and intellectual value have stimulated a number of follow-up works. Particularly for $N = 2$, it was theoretically proven that not only faint two-photon light, but also stronger two-mode squeezed light can be used for this purpose [87]. On the experimental side, we would like to acknowledge the success in driving coherent [88] and incoherent [82] two-photon processes with SPDC light.

Discerning the objects' features with resolution exceeding the Rayleigh diffraction limit is possible in setups similar to two-photon *Ghost imaging* setup [89] but relying on multi-photon entangled states such as GHZ or W states [41]. Alternatively, axial resolution can be enhanced by a factor of two realizing a quantum version of optical coherence tomography measurement with two-photon light [90]. In this case one makes use of the signal-idler intensity correlation time being much shorter than their individual coherence times.

The resolution of small lateral displacement measurements is limited by the shot noise to the value

$$d_0 \approx \sqrt{\frac{\pi}{8N}} w_0, \quad (3.7)$$

where w_0 is the Gaussian width of a TEM_{00} probe beam focused onto a split-field detector, and N is the number of detected photons. It has been shown [91] that by composing the probe beam out of coherent and squeezed optical beams as shown in Fig. 3.4, the shot-noise resolution limit (3.7) can be improved by approximately a factor of two.

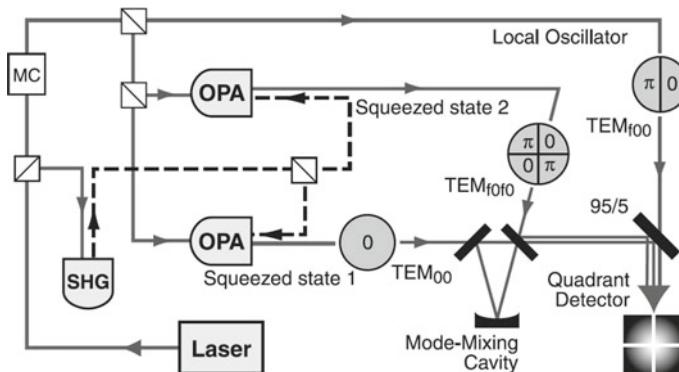


Fig. 3.4 A superposition of squeezed light and coherent local oscillator with segmented phase shifts enables the beam displacement measurement with precision exceeding the standard quantum limit. SHG is second harmonic generator, OPA is optical parametric amplifier, MC is mode cleaner, 95/5 is a beamsplitter with 95% reflectivity. TEM_{f0f0} and TEM_{f00} designate a formerly TEM_{00} mode modified by split phase plates. Dashed lines show 532 nm light; solid lines show 1064 nm light. Reprinted from [91]

Quantum Information Processing

The concept of quantum information processing, or quantum computing, was conceived in 1982 by Richard Feynman [92]. At the heart of this concept is a notion that a quantum superposition principle can be utilized to implement a large number of computations in parallel. To implement such quantum parallelism, logic operations of a quantum computer must be performed by quantum systems. It should be noted, however, that in order to access the results of this parallel computation one has to perform a measurement which is equivalent to a projection onto just one result. Therefore, one benefits from this parallelism only if the single measurement already provides an advantage, such as in the Shor algorithm [93] where a quantum interference phenomenon is utilized to find prime factors of a large number faster than it is possible by the classical search. Note that the *classical* optical interference can be used in a similar way [94].

Instead of encoding information in bits that take on binary values 0 or 1, these systems encode it in *qubits*, allowing any superposition of the binary values. A qubit may be implemented in various two-level physical systems, such as an atom, ion, spin-1/2 particle, and many others. To distinguish such systems from photons, we will call them *massive*. Polarization of a photon, as well as its localization in two spatial or frequency modes, also can be used as a qubit. The advantage of optical qubits over massive ones is slow decoherence of the former: photons hardly interact with ambient electromagnetic or gravity fields.

This advantage however turns into a disadvantage when it comes to implementation of quantum logical operations that require photon-photon interaction. Such interaction can be facilitated using optical nonlinearity at the single photon level. Several approaches to building quantum gates based on nonlinear response of optical media have been theoretically discussed. One of these approaches is the Quantum Zeno Blocade which can be realized based on two-photon absorption [95], electromagnetically induced transparency [96, 97], or on the second-order polarizability of optical nonlinear crystals enhanced by high-*Q* cavities [98–100]. Several experimental demonstrations of these techniques have been performed with multi-photon (typically, weak coherent) states [101, 102], however functional photonic quantum gates so far remain beyond the reach.

This difficulty has led to the concept of quantum network [103], where transmission of information is performed by photonic qubits, while its processing is performed by massive qubits. Various types of massive qubits have been successfully coupled to single photons, including atoms [104–111] and quantum dots [112]. Nitrogen vacancy centers in diamonds have been also proposed for this application [113, 114].

Building a quantum network requires non-classical light sources whose central wavelength and optical bandwidth are compatible with the massive qubits. In the most straightforward way this can be achieved by using the same atomic transition for the generation of non-classical light (see discussion in Sect. 3.2.1), and then for transferring quantum information to atomic qubits [115–118]. Alternatively, narrow-line parametric light sources discussed in Sect. 3.2.2 can be used. Note that while generating narrow-band squeezed light or squeezed vacuum is relatively easy by

operating an OPO source above the threshold, generation of equally narrow-band photon pairs below the threshold is more difficult, as it requires tunable resonators with very high Q -factor [119–123]. For many quantum information applications such sources also need to be strictly single-mode, which has been recently achieved using whispering gallery mode (WGM) [124] and waveguide [125] resonators.

Using massive qubits often requires low temperatures, very low pressure vacuum, thorough shielding of ambient fields, and entails other serious technical complications. The concept of *linear* quantum computing [126] strives to avoid these complications. There are no massive qubits in a linear quantum computer, but there are also no photon-photon interactions. This interaction is replaced by a measurement process followed by feed-forward to or post-selection of the remaining photons. This procedure is certainly nonlinear (and even non-unitary), and can be used to implement quantum logic operations over a sub-space of a larger Hilbert space.

In higher-dimensional Hilbert spaces photonic *qutrits* [127, 128] and even *ququarts* [129, 130] can be introduced as useful notions. As an example, a photon qutrit encoded in polarization has three basis states: $|\uparrow\downarrow\rangle$, $(|\leftrightarrow\downarrow\rangle + |\downarrow\leftrightarrow\rangle)/\sqrt{2}$, and $|\leftrightarrow\leftrightarrow\rangle$. A ququart basis consists of four states and can be easily envisioned if we further lift the frequency degeneracy, or couple the photon pair into different spatial modes. Usually these states are discussed in the context of quantum secure communications using alphabets with higher than binary basis.

Transmission of information by photonic qubits presents sufficient interest by itself, besides being a quantum computer building block. The fundamental property of a qubit is that it cannot be cloned, or duplicated. Such cloning would be incompatible with the linearity of quantum mechanics [131]. Therefore, the information encoded in qubits can be read only once; in other words, it cannot be covertly intercepted. This property of qubits served as a foundation for the original quantum key distribution (QKD) protocol BB84 [132], and for numerous and diverse QKD protocols that emerged later. QKD is the least demanding application of non-classical light reviewed in this chapter, and the only quantum optics application known to us that has been relatively broadly commercialized to-date. Discrete variables QKD can be successfully implemented even with weak coherent light, e.g. strongly attenuated laser pulses, which adequately approximate single-photon states. Similarly, non-orthogonal coherent states of light can be successfully used in continuous variables QKD [133].

Coherent states are pure quantum states unlike thermal states and thus qualify as non-classical states (see Appendix for discussion). For some quantum protocols coherent states suffice, for others they do not. Furthermore, it is often argued that much of their properties can be described by classical models. For all these reasons we concentrate the discussion on states which are more non-classical than coherent states.

Some proposed quantum information protocols relying on non-classical light fall between the QKD and quantum computing in terms of architecture and complexity. One of such protocols is the *quantum commitment*. It is designed to allow Alice to fix (“commit”) an observable value in such a way that Bob cannot learn this value until Alice reveals it. Alice, on the other hand, cannot change her commitment after it has

been made. Originally proposed in 1988 [134], this protocol has been experimentally demonstrated [135] in 2013 with an added benefit of closing a loophole present in the original proposal. Other protocols proposed for implementing quantum secret sharing among multiple parties [133] and quantum digital signatures [136] may be used in the context of quantum money, quantum voting, and other visionary applications. In the following section we review the sources of non-classical light, which is the main objective of this chapter.

3.2 Sources of Non-classical Light

3.2.1 Atoms Real and Artificial

Atoms

The early interest in non-classical, and in particular entangled, optical states was stimulated by the quest for experimental violations of Bell’s inequalities. The first successful and statistically reliable violation was reported in 1972 by Freedman and Clauser [137]. They used a cascade two-photon transition in calcium beam producing a polarization-entangled pair of blue and green photons, and performed a polarization-based Bell measurement which has shown a six standard deviations violation. Therefore the conceptually more advanced two-photon entanglement was observed with atomic sources prior to a more straightforward antibunching effect.

Photon antibunching in resonance fluorescence from a coherently driven two-level atom is easy to understand. Once the atom emits a photon, it occupies the ground state and cannot emit another photon for a period of time of the order of the excited state lifetime (in the weak excitation regime), until the interaction dynamics drives the atom back to the excited state. Hence the Poissonian statistics of the coherent pump photons is converted to sub-Poissonian statistics of fluorescence photons, leading to a state whose photon-number fluctuation is reduced below the shot noise limit typical for coherent light, such as shown in Fig. 3.1f.

Antibunching in resonance atomic fluorescence was predicted back in 1976 by Carmichael and Walls and observed in 1977–78 by two different research groups using beams of sodium atoms, see the review [3] for details. More recently, four-wave mixing in a rubidium vapor cell was used to produce and characterize heralded Fock-basis qubits $\alpha|0\rangle + \beta|1\rangle$ [138].

A sodium atomic beam passing through an optical cavity was also used for the first demonstration of squeezed light in 1985 [139]. Soon after that the first magneto-optical traps were implemented. They allowed to suppress the thermal motion of atoms and—associated with it—the dephasing, which increased the observed squeezing from 0.3 dB [139] to 2.2 dB [140]. Even stronger was the two-mode squeezing observed by seeding one [141] or both [142] of these modes with weak coherent light. In these experiments the squeezing was measured to be 3.5 dB (8.1 dB corrected for losses) and 3 dB (over 3.5 dB corrected for losses), respectively. This technique has

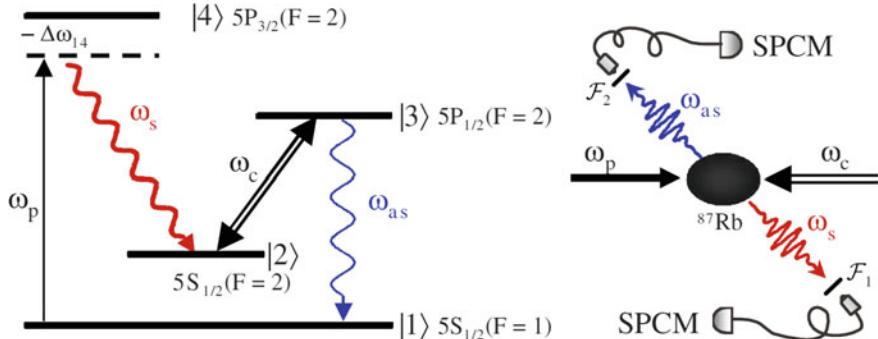


Fig. 3.5 A double- Λ configuration of ^{87}Rb transitions involved in the four-wave mixing process generating non-classical light and the experimental diagram. Reprinted from [144]

a potential for tailoring the spatial structure of multimode non-classical light, e.g. generating twin beams carrying orbital angular momentum [143].

A single pump laser was used in experiments [3, 139–143]. To suppress the effect of thermal motion, the four-wave mixing process can be driven by two *different*, counterpropagating, laser beams in a configuration typical for saturation absorption spectroscopy. This technique has allowed for generation of very high flux of photon pairs with controllable waveform, see [144] and references therein. Such pairs can be used for heralded preparation of nearly single-photon pulses. Moreover, the ground-state coherence in cold atomic ensembles is sufficiently long-lived to allow the “read” laser pulse to arrive with a substantial delay after the “write” pulse, which allows one to control the delay between the emitted heralding and the heralded photons [145–147]. A controlled delay is in fact just a special case of temporal shaping of the biphoton correlation function, which can be achieved with the “read” pulse profile manipulation [148].

Quantum optics researchers favored alkali atomic gases because of their strong resonant Kerr response. A typical energy diagram of this process, called a double- Λ configuration, is shown in Fig. 3.5. This diagram is drawn specifically for ^{87}Rb D1 and D2 manifolds, but its analogues can be realized in various atomic species. Strong pump and control optical fields have frequencies ω_p and ω_c , corresponding to D2 and D1 transition wavelengths, respectively. Generated quantum (two mode squeezed) light has the Stokes and anti-Stokes frequencies ω_s and ω_{as} , respectively. The energy and momentum conservation requires $\omega_p + \omega_c \approx \omega_s + \omega_{as}$ and $\mathbf{k}_p + \mathbf{k}_c \approx \mathbf{k}_s + \mathbf{k}_{as}$, where the approximations arise from neglecting the momentum recoil and kinetic energy that maybe carried away by the atom. Note that the momentum conservation allows for a very broad angular spectrum of the emitted light in the case of counter-propagating ($\mathbf{k}_p + \mathbf{k}_c \approx 0$) beams.

Another important feature of atomic Kerr media is that its response may be sensitive to light polarization. This can lead to nonlinear phenomena such as polarization self-rotation (see [149] and references therein), where one polarization is amplified

while the orthogonal polarization is deamplified. If the input light is polarized linearly [149, 150] or circularly [151], the vacuum field in the orthogonal polarization becomes squeezed.

Coupling atomic media with optical cavities opens up the field of cavity quantum electrodynamics (cQED), rich with non-classical phenomena. Even a single atom strongly interacting with an optical mode can generate squeezed light [106]. It can also be used to implement a photonic blockade [152], leading to a photon turnstile capable of generating single photons on demand [104, 153]. In terms of quantum systems engineering, this can be considered as a next step after delayed heralded single-photon generation, and two steps after single photons generated at random times. A real or artificial atom strongly coupled to a cavity mode is also predicted to be capable of generating the “N-photon bundles” [154], arguably equivalent to flying Fock states [155].

Once generated, the non-classical states need to be routed in a decoherence-free manner towards the information-processing nodes or to detectors. The single photon routing controlled by other single-photon states would enable quantum logic operations on photons, and make an optical quantum computer possible. Serious efforts have been made in this direction. An optical transistor was reported [107], in which a single control photon induced a ground-state coherence in a cold Cesium cloud, affecting the transmission of a delayed probe pulse. In a more recent work [109], a single-photon switch based on a single Rubidium atom interacting with the evanescent field of a fused silica microsphere resonator was demonstrated. This system was shown capable of switching from a high reflection (65%) to a high transmission (90%) state triggered by as few as three control photons on average (1.5 photons, if correction for linear losses is made).

Finally, let us point out that single molecules can be similarly used as sources of single photons, as demonstrated by a significant measured antibunching [156, 157]. Molecular sources can operate at room temperature in the on-demand mode [158].

Artificial Atoms

Discrete level spectra are available not only in atoms but also in solid-state nanosystems, such as quantum dots, carbon nanotubes, nitrogen vacancy (NV) centers in diamond, or impurities in semiconductors. Because of this property such systems are often referred to as “artificial atoms”. They too have been actively utilized as sources of non-classical light. The physical mechanism regulating the photons statistics of an artificial atom emission is very similar to that of real atoms.

While an optical photon absorption by an atom causes an electron transition from the ground to an excited state, in quantum dot it causes generation of an electron-hole pair, called an exciton. The recombination of this exciton is responsible for the resonance fluorescence of a quantum dot. Applications of this process for single-photon sources are reviewed in [159, 160]. Such sources often require liquid helium cooling, although the first demonstration of non-classical light emitted from a quantum dot was done in year 2000 by Michler et al. at room temperature [161]. In this experiment a single CdSe/ZnS quantum dot was driven by a resonant constant wave (CW) pump laser. Its fluorescence had sub-Poissonian photon-number distribution

with $g^{(2)}(0) = 0.47 \pm 0.02$. More recent quantum dot based sources [162, 163] also can operate at room temperature exhibiting non-classical anti-bunched photon statistics in pulsed regime, although their antibunching is significantly stronger at liquid helium temperatures.

In carbon nanotubes, two-photon generation is suppressed due to Auger processes and excitons localization. Antibunching of the light emitted by such systems can be very strong, reaching the value of $g^{(2)}(0) = 0.03$ at 4.2 K [164].

In contrast with quantum dots and carbon nanotubes, NV centers in diamond provide the most stable quantum emitters at room temperature. In [165], a CW emission from a single NV center in a diamond nanocrystal was coupled to a $4.84\text{ }\mu\text{m}$ in diameter polystyrene microspherical resonator. The non-classical character of the single quantum emitter was verified by measuring $g^{(2)}(0) \approx 0.3$, while the coupling to the WGMs was evident from a discrete spectrum of the emission. NV-center based pulsed single-photon sources also operate at room temperature reaching nearly the same antibunching figure [166].

The power fluctuation measurements carried out in [161–163, 166] allowed to probe the Wigner function only in the radial direction (c.f. Fig. 3.1). A more advanced measurement also providing the access to the orthogonal quadratures was carried out by Schulte et al. who used a local oscillator with variable phase in a heterodyne setup [167]. They also studied the amount of squeezing as function of the excitation power.

Just as with real atoms, coupling quantum dots to microcavities provides access to the benefits of cQED. One of these benefits is the improved collection efficiency. Because of the high Purcell factor of the microcavities, a quantum dot fluorescence is preferentially radiated into the cavity modes and can be conveniently collected. Press et al. [168] have been using micro-pillar structures for this purpose. A micro-pillar resonator shown in Fig. 3.6a measures about a micron in diameter and five micron tall. It is complete with Bragg mirrors at both ends, each consisting of approximately 30 pairs of AlAs/GaAs layers. A layer of InGaAs quantum dots is grown in the central anti-node of the cavity. The structure is cooled to 6–40 K and pumped by a pulsed mode-locked laser. Photons collected from the cavity were antibunched with $g^{(2)}(0) \approx 0.18$. A different design shown in Fig. 3.6b uses a layered structure where a pillar cavity is defined by cutting trenches of various shapes [169]. This shape allows one to control the polarization dispersion of the resonator and to generate single photons with a desired polarization. Quantum dots have been coupled not only to pillar or planar cavities, but also to WGM resonators. For example, strong coupling regime was achieved with a single GaAs [170] and InAs [112, 171] quantum dots.

Instead of a cavity, a quantum dot can be coupled to a single-mode on-chip waveguide [172]. This approach not only allows to generate strongly non-classical ($g^{(2)}(0) < 0.1$) light, but also leverages scalable on-chip photonic technology. Operating these systems in pulsed mode gives them a much desired “single photon on demand” quality.

Quantum dots can support not only single excitons, but also biexcitons. Recombination of a biexciton leads to emission of a photon pair, similarly to a two-photon emission from an atom in a Freedman and Clauser experiment. This process can go through two different intermediate states, realizing two quantum-mechanical paths

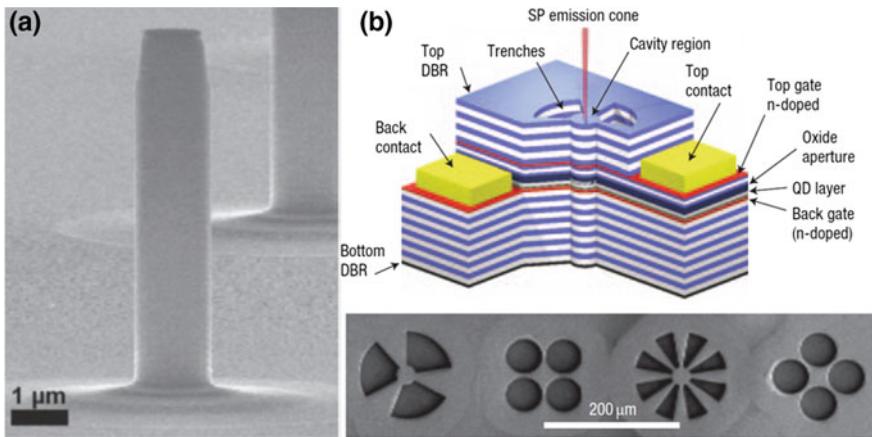


Fig. 3.6 A micro-pillar optical resonator (a) has Bragg mirrors at the base and on the top, providing strong coupling of quantum dots (embedded near its center) to a vertical mode. In a different design (b) the resonator is formed by cutting trenches of various shapes (shown in the bottom) in a layered structure. Reprinted from [168, 169]

for biexcitonic recombination. In experiments [173–176] the photon pair emitted along one path is polarized vertically; along the other, horizontally. Thus recombination of such a biexciton creates an optical Bell state $|\Phi^{(\pm)}\rangle$ introduced in (3.4), provided that the polarization terms are not “tagged” by either the final (ground) state of the quantum dot, or the optical wavelength. Then violation of Bells inequality is possible, and in fact has been observed with a confidence of five standard deviations [175].

The biexciton recombination process is broadband enough to provide a significant wavelength overlap even if perfect wavelength degeneracy cannot be achieved. This allows one to erase the wavelength distinguishability by spectral filtering and achieve a polarization-entangled state capable of violating Bells inequality by more than three standard deviations [173]. Similar mechanisms can lead to polarization-entangled photon pairs emission from impurities in semiconductor [177] and from the hybrid piezoelectric-semiconductor quantum dot systems [178]. The latter system has been also used to demonstrate Bells inequality violation.

Generation of entangled photon pairs by quantum dots is unique in that the pairs themselves have sub-Poissonian statistics, which allows to generate single photon pairs using a pulsed pump. This aspect of the quantum dot entangled light sources was highlighted by Young et al. [179], who demonstrated the triggered emission of polarization-entangled photon pairs from the biexciton cascade of a single InAs quantum dot embedded in a GaAs/AlAs planar microcavity. They also showed that quantum dot engineering can reduce the energy gap between the intermediate states, minimizing or removing the need for thorough spectral filtering. Deterministically exciting biexcitons by optical π -pulse, Müller et al. [180] have demonstrated a true “polarization entangled photon pair on demand” operation with unprecedented anti-bunching parameter $g^{(2)}(0) < 0.004$ and high entanglement fidelity (0.81 ± 0.02).

3.2.2 Parametric Down Conversion

Spontaneous parametric down conversion (SPDC), optical parametric amplification (OPA) and oscillation (OPO) are among the most important sources of non-classical light. All these closely related processes are enabled by the second-order nonlinear response of non-centrosymmetric optical crystals, characterized by quadratic susceptibility $\chi^{(2)}$. This process, originally called parametric scattering or parametric fluorescence, was first observed in 1965 [181] and widely studied later. From the quantum point of view, i.e. in terms of photon pair emission, this process was first discussed in 1969 by Klyshko [182]. One year later, the “simultaneity” of these photons (called the *signal* and *idler*) was observed by Burnham and Weinberg [61].

We now know that the reported “simultaneity” reflected the resolution of the time measurements rather than the physical nature of the biphoton wavefunction. The signal-idler correlation time is finite, and is closely related to their optical spectra and the group velocity dispersion (GVD) of the parametric nonlinear crystal [183]. The temporal correlation function can take on different forms for different types of phase matching [184–187], with the width ranging over six orders of magnitude: from 14 fs for free-space SPDC in a very short crystal [185] to 10–30 ns for SPDC in a high-finesse optical resonators [119–122, 124, 188–190]. Shaping this correlation function is an important problem in quantum communications. With SPDC lacking the photon-storage capability available to the atomic sources, this problem is quite challenging. One possible approach is by interferometric tailoring of the SPDC spectra using two or possibly more coherently pumped crystals [84, 191, 192]. Another approach is based on using a dispersive media [186]. There is also a proposal for leveraging the *temporal ghost imaging* [193], which is similar to spatial ghost imaging [89] but relies on temporal rather than spatial masks (implemented e.g. by electro-optical modulators) [194, 195].

Parametric down conversion has been described in great detail in many books and papers, which spares us the necessity to reproduce all the analysis and derivations here. Let us just list the most fundamental facts. The energy and momentum conservation for the pump, signal and idler photons impose the phase matching conditions

$$\omega_p = \omega_s + \omega_i, \quad (3.8)$$

$$\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i, \quad (3.9)$$

where the frequencies are related to the wave numbers by the dispersion relations $\omega = ckn(\lambda)$. It is the combination of these three constraints that is responsible for the free-space SPDC light appearing as a set of colorful rings. In most materials, normal chromatic dispersion of the refractive index $n(\lambda)$ prohibits parametric phase matching by making (3.8) and (3.9) incompatible. However it can be compensated by polarization dispersion in birefringent materials. For example, the pump polarization can be made orthogonal to both signal and idler polarizations, which is known as Type-I PDC configuration. Alternatively, either signal or idler polarization can be parallel to that of the pump in Type-II PDC. Type-0 PDC, when all three fields

are polarized in the same plane, can be attained by using various periodical poling techniques. Periodical poling modifies (3.9) by adding or subtracting a multiple of the poling structure wave vector $\mathbf{e} 2\pi/\Lambda$, where Λ is the poling period and \mathbf{e} is its direction.

A pair of coupled signal and idler modes with photon annihilation operators a_s and a_i , respectively, is governed by the evolution operator

$$\hat{U}(t) = \exp \left\{ -\frac{i}{\hbar} \int_0^t \hat{H}_{\text{int}}(t) dt \right\}, \quad \hat{H}_{\text{int}}(t) = i\hbar g(t)(a_s^\dagger a_i^\dagger - a_s a_i). \quad (3.10)$$

This is an approximation assuming that the pump field can be treated classically, i.e. that one can neglect the annihilation of one pump photon for every creation of a signal-idler photon pair. The function $g(t)$ in (3.10) describes parametric interaction:

$$g = 2\pi\sigma(t)(\chi^{(2)} : \mathbf{e}_p \mathbf{e}_s \mathbf{e}_i) \frac{\sqrt{\omega_s \omega_i}}{n_s n_i}, \quad (3.11)$$

where $(\chi^{(2)} : \mathbf{e}_p \mathbf{e}_s \mathbf{e}_i)$ is the scalar product of the nonlinear susceptibility tensor with the interacting fields unit vectors. The overlap integral

$$\sigma(t) = \int \psi_s(\mathbf{r}) \psi_i(\mathbf{r}) E_p^*(\mathbf{r}, t) dV \quad (3.12)$$

is calculated for the normalized modes eigenfunctions $||\psi(\mathbf{r})_{s,i}|| = 1$ and the pump field envelope $E_p(\mathbf{r}, t)$. This integral enforces the momentum conservation (3.9) for the plane-wave modes.

The time integral $G \equiv \int_0^T g(t) dt$ is called the parametric gain. Here the interaction time T is determined by the crystal length. The effective interaction length however can be shorter than the crystal length for short pump pulses, when significant longitudinal walk-off between the pump and parametric pulses occurs due to the GVD. Note that depending on the pump phase G may take on negative values, leading to de-amplification.

Spontaneous Parametric Down Conversion

Vacuum-seeded parametric down conversion, or SPDC, is probably the most widely used non-classical light source made famous by Bell's inequality violations, early QKD demonstrations, quantum teleportation, and a number of other remarkable achievements PDC has made possible. This process has been realized in low and high gain regimes, in free space, in single transverse mode waveguides, and in optical resonators. In the low-gain regime, this process is adequately described by expanding the evolution operator \hat{U} from (3.10) into a power series. The leading term of the expansion represents a two-mode vacuum, next term is a signal-idler photon pair, the third term represents two such pairs, and so on. The amplitudes of these terms

form the same power series as for a thermally populated mode [1, 70, 196], which determines the peak value of the Glauber correlation function for a weakly populated SPDC mode: $g^{(2)}(0) = 2$. This also allows one to introduce the *effective temperature* for SPDC emission [70].

Free-space SPDC provides for a multimode source of spatially-entangled biphotons, which can be used in two-photon imaging discussed in Sect. 3.1.3. This type of entanglement arises from the momentum conservation (3.9). Indeed, even with strictly constrained (e.g., by band-pass filters) optical wavelengths, there are many indistinguishable ways the transverse momentum conservation $k_s^\perp + k_i^\perp = 0$ can be achieved. On the other hand, selecting a single transverse signal mode, and the paired with it idler mode, one obtains a frequency-entangled state (3.5).

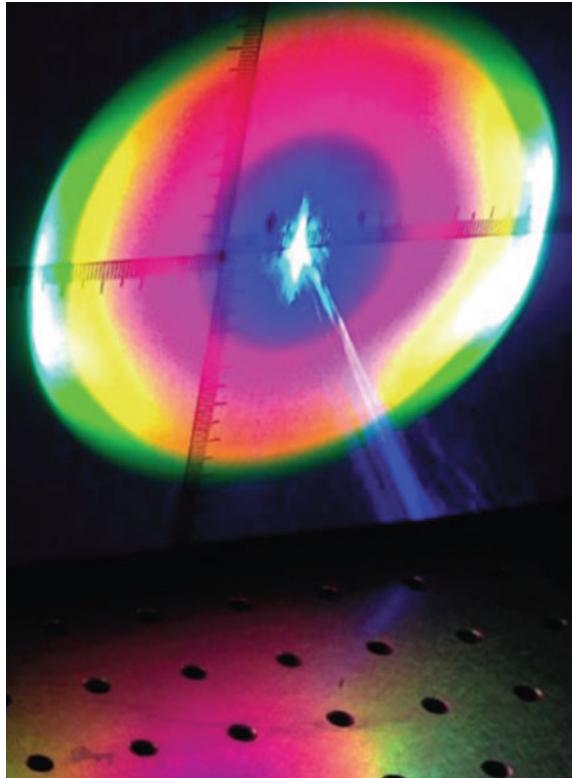
Type-II SPDC offers an interesting configuration [197] wherein the same pair of spatial modes A and B can be populated by orthogonally polarized signal and idler in both possible permutations, leading to a polarization-entangled state such as $|\Psi^{(+)}\rangle_{AB}$ in (3.4). A closer look shows that this state is also frequency-entangled as in (3.5). Such states that are entangled in more than one degree of freedom at once are called hyperentangled [198].

Polarization entangled photon pairs can also be generated in Type-I SPDC, in a clever configuration of two crystals whose optical axis planes are perpendicular to each other [197]. This configuration provides even more flexibility than the polarization entanglement generation in a Type-II crystal: by varying the phase between the pump field projection on the two crystals' axes (e.g., varying the pump polarization ellipticity), as well as manipulating the polarization and phase of the signal and idler photons between the crystals, one can generate any polarization-entangled state in Hilbert space spanned by the Bell-states basis (3.4), as well as some of mixed states [197].

Parameters of SPDC biphoton sources such as their wavelengths, bandwidth and pair production rate may vary considerably. Because of accidental generation of multiple photon pairs, ultra high pair rate associated with large G is not always desirable. It is often more important to minimize the chance of accidentally generating a second pair during the measurement. In the limit of very fast measurements it is also important to generate sufficiently few (much less than one on average) photons per *coherence time*, i.e. per longitudinal mode. If this number exceeds unity, then $G > 1$ as well, and the power series expansion of the evolution operator \hat{U} does not converge. This means that the already generated parametric photons make a stronger contribution to the further PDC process than the vacuum photons, i.e. we enter the regime of parametric super luminescence. This is accompanied by a transition from thermal (Gaussian) photon number statistics to Poissonian statistics, typical for laser light.

However the parametric light remains non-classical even in the high-gain regime. When the signal and idler are distinguishable, the light is two-mode squeezed, which can be established by measuring the photocurrents difference in the signal and the idler detectors and finding it below the shot noise level. When the signal and idler are indistinguishable, we have the squeezed vacuum state such as shown in Fig. 3.1d, whose photon-number basis expansion consists of only even terms and

Fig. 3.7 It is incorrect to think of non-classical light as always faint. Bright parametric light on this photo is a macroscopic quantum state. Courtesy of M. V. Chekhova



$\langle p \rangle = \langle q \rangle = 0$. Let us recall that if the signal and idler have the same frequency and the distinguishability is only based on polarization or emission direction, a conversion between two-mode squeezing and squeezed vacuum is trivially accomplished with a polarizing or non-polarizing beam splitter, respectively. In these cases the terms “two-mode squeezing”, “squeezed vacuum” and even “two-mode squeezed vacuum” are often used interchangeably.

Parametric gain determines the mean photon number in a mode $\langle N \rangle = \sinh^2(G)$ as well as the squeezing parameters: $q_{\text{out}} = q_{\text{in}} \exp(G)$, $p_{\text{out}} = p_{\text{in}} \exp(-G)$. We have used these relations calculating the Wigner function shape in Fig. 3.1d. Even in strongly pumped parametric processes, G is typically less than ten. The record value of $G \approx 16$ is reported in [199]. But let us not be deceived by these small numbers. Unlike a gain of a common amplifier, parametric gain is exponential, see (3.10), so SPDC with $G \approx 16$ produces over 10^{13} photons per mode. Therefore multimode light generated in parametric down conversion can be quite strong in terms of the optical power, see Fig. 3.7, but still non-classical.

Multimode SPDC light is useful for imaging and similar applications. Here the number of modes can be compared to the number of pixels, and directly translates to the spatial resolution. Single mode SPDC, on the other hand, is often desirable

for quantum communication applications, when the presence of multiple mutually incoherent modes is equivalent to the loss of the phase information, or decoherence. Spatial and frequency filtering can be employed to purify the SPDC mode structure, but this approach is not power-efficient if the initial source has too many modes. The number of excited transverse modes can be reduced, even to one, by using waveguides instead of bulk crystals. This provides a dramatic benefit over the filtering approach in terms of useful photon pair rate. For example, it is possible to generate and collect about 100,000 photon pairs per second with only 0.1 mW pump [200].

The number of frequency or temporal modes can be controlled by matching the SPDC linewidth, determined by the source length, geometry and GVD, with the transform-limited spectrum of the pump pulse. This can be done e.g. by adjusting the pump pulse duration.

Combining these two techniques, nearly single-mode parametric sources can be realized [201]. Let us also mention that the birefringent properties of parametric crystals can make the gain so selective that in the super luminescence regime even free-space parametric sources can approach single-mode operation [202].

Multipartite multiphoton states can be prepared in SPDC process by combining two or more identical coherently pumped sources [203, 204], or by splitting multiphoton states from a single source [38, 40, 205]. These experiments are difficult because of the thermal statistics of SPDC pairs. Although higher photon-number states are less likely to emerge, they are more likely to cause a detection event with imperfect ($\eta < 1$) detectors. Suppressing such events requires limiting the overall photon flux, which leads to very low data rate, typically of the order of 1/s for four-photon measurements and 1/h for six-photon measurements.

Optical Parametric Amplification

If a degenerate or non-degenerate parametric process has non-vacuum inputs in the signal and idler modes, it may amplify or de-amplify the input beams depending on the relation between the sum of their phases and the phase of the pump, which determines the sign of G . If one of the inputs, e.g. the idler, is in vacuum state for which the phase is not defined, then the signal will always be amplified. On the phase space diagram it will appear as both displacement and quadrature-squeezing [29].

Like SPDC, OPA is a common technique for generating non-classical light. This technique is most suitable for squeezing coherent light pulses seeding the OPA. A 2 dB [206] and then 5.8 dB [207] squeezing of 270 ns long pulses in a degenerate Type-II parametric amplifier was demonstrated. A thousand times shorter squeezed vacuum pulses (250 fs, 1.7 dB squeezing) were generated in a Sagnac interferometer configuration using periodically poled lithium niobate crystal [208].

Continuous wave coherent states can also be used for seeding the OPAs, which allows for precise control of the local oscillator phase. This technique has been used to generate quadrature-squeezed light by injecting fundamental laser light into a degenerate OPA waveguide made from periodically poled KTP and pumped by the second harmonic of the fundamental laser light, reaching 2.2 dB of squeezing [209]. Realizing a similar process in a monolithic cavity with highly reflective coating on the parametric crystal facets, 6 dB of squeezing has been reached [210]. Using a Type-II

OPA in a bow-tie cavity yielded 3.6 dB of polarization squeezing [211], which corresponds to reduction of the quantum uncertainty of the observables associated with the Stokes operators (3.6).

Often the OPA seed signal itself is generated in another SPDC process taking place in a similar crystal and pumped by the same pump. This configuration is sometimes called a nonlinear interferometer. We have already encountered it discussing the spectroscopy applications in Sect. 3.1.3. The high mode selectivity of such interferometers has allowed to implement a nearly single-mode squeezed vacuum source without a significant decrease in the output brightness [192, 212]. It is also possible to cascade more than two OPAs. A system of three OPAs reported in [213] has boosted the two-mode squeezing from 5.3 dB measured after the first OPA to 8.1 dB after the third one.

Parametric Processes in Cavities

An amplifier can be turned into an oscillator by providing a positive feedback, e.g. by placing the amplifying media into an optical cavity. Such a setup was used in the first demonstration of parametric squeezing in 1986 by Wu et al. [214]. In this experiment, frequency-doubled 1064 nm laser light pumped a degenerate OPO system consisting of a lithium niobate crystal inside a Fabri-Perot resonator. The same fundamental laser light was used as a local oscillator in homodyne detection of the squeezed vacuum. 3.5 dB of squeezing was measured. In 1992 this result was slightly improved to 3.8 dB with a bow-tie cavity [75]. This configuration was further improved by using periodically poled KTP crystals, which reduced the linear and pump-induced absorption and eliminated the transverse walk-off. 7.2 dB of squeezing was demonstrated in 2006 [215], and 9 dB in 2007 [216]. Thorough stabilization of a cavity allowed generation of a narrow-band, 5 dB squeezed vacuum matching the rubidium D1 line [217]. Using a monolithic cavity boosted the squeezing to 10 dB in 2008 [218] and to 12.7 dB in 2010 [219]. Most recently, a new record, 15 dB of squeezing, was reported [68].

The experiments [68, 75, 139, 214–219] were carried out below the OPO threshold. This means that the mean photon number per mode was below unity, or in other words, the process was predominantly vacuum-seeded, in contrast to the case of self-sustained oscillations. In this sense a sub-threshold OPO can be compared to a very long crystal in an SPDC experiment. By contrast, operating an OPO *above* the threshold turns it into a laser. This is not an ordinary laser, however. A non-degenerate OPO laser emits two beams that are quantum-correlated, or two-mode squeezed. This has been demonstrated already in 1987 by Heidmann et al. who used a Type-II OPO to generate a few milliwatts in each near-degenerate signal and idler beams [220]. In a few years the same approach yielded 8.5 dB of two-mode squeezing [221], which had remained a squeezing world record perhaps for the longest time.

The photon number correlation between the signal and idler beams can be used to prepare sub-Poissonian light in either one of these beams. This was demonstrated in 1988 by Tapster et al. [222] who detected the fluctuations of the signal beam power emitted in a Type-I SPDC process from a KDP crystal in frequency-degenerate but non-collinear configuration, and fed them back to the pump power thereby achieving

the photon-number squeezing in the idler beam. A variation of this experiment was performed later with a sub-threshold OPO [223], in which case the signal measurement was *fed forward* to a fast intensity modulator placed in the idler beam. In [223] one can also find an extensive theoretical analysis of both feedforward and feedback techniques applied for preparing sub-Poissonian light in PDC. In Sect. 3.2.4 we will see how both these techniques can be applied to other types of lasers to generate non-classical light.

This approach received an interesting development in 2003 [224], when instead of actively using the signal power fluctuations in a feedback or feedforward loops, Laurat et al. used them for conditioning the signal-idler squeezing measurement. Only those measurements were retained when the fluctuations were below a certain threshold. Thereby a continuous-variables post-selective measurement was implemented, which allowed to observe 7.5 dB of squeezing.

Discussing the quantum information applications of non-classical light, we mentioned the importance of making the source narrow-band enough to match the optical transitions widths in gas phase ensemble quantum memories, often implemented with atoms or ions. An OPO provides such an opportunity. Above the threshold, its line can be considerably narrower than the cold cavity linewidth due to the Schawlow-Townes effect. Thus even with modest cavities OPO light can match the narrow atomic transitions. Hald et al. used this approach to observe spin squeezing of cold atomic ensemble induced by interaction with squeezed vacuum [225]. Later it was shown that such a spin-squeezed atomic state can regenerate the squeezed vacuum, thereby verifying its storage [226].

It is more difficult to achieve narrow-line OPO operation below the threshold. Usually it requires external high- Q filter cavities [119, 120, 122, 227] or post-selection [228] techniques that considerably reduce the signal rate, as well as introduce inevitable losses at the edges of the filter windows. It would be desirable to generate photon pairs directly into a single or a few easily separated modes. This became possible by using WGM micro-resonators.

In WGM resonators light is guided along a smooth optical surface of rotation by continuous total internal reflection, similarly to how sound is guided in their name-sake acoustical analogues. WGM resonators defy the postulate of light propagating in a straight line in the most profound way: here the light ray bends at every point. The WGM eigenfunctions inside of a spherical resonator are

$$E_{mlq}(r, \theta, \varphi) = E_0 j_m(nk_q r) P_l^m(\cos \theta) e^{im\varphi}, \quad (3.13)$$

where (r, θ, φ) are usual spherical coordinates, j_m is the spherical Bessel function of order m , P_l^m are the associated Legendre Polynomials, and E_0 is the amplitude. The eigenvalue k_q for a given radial mode number $q = 1, 2, 3, \dots$ is found by matching the internal Bessel and external Hankel eigenfunctions according to the boundary condition at the resonator rim $r = R$. For relatively large WGM resonators with small evanescent field the approximation $j_m(nk_q R) = 0$ yields quite accurate results.

It is convenient to introduce $p = l - |m| = 0, 1, 2, \dots$ which gives the mode order in the θ direction, similarly to how q gives it in the radial direction. Intensity distri-

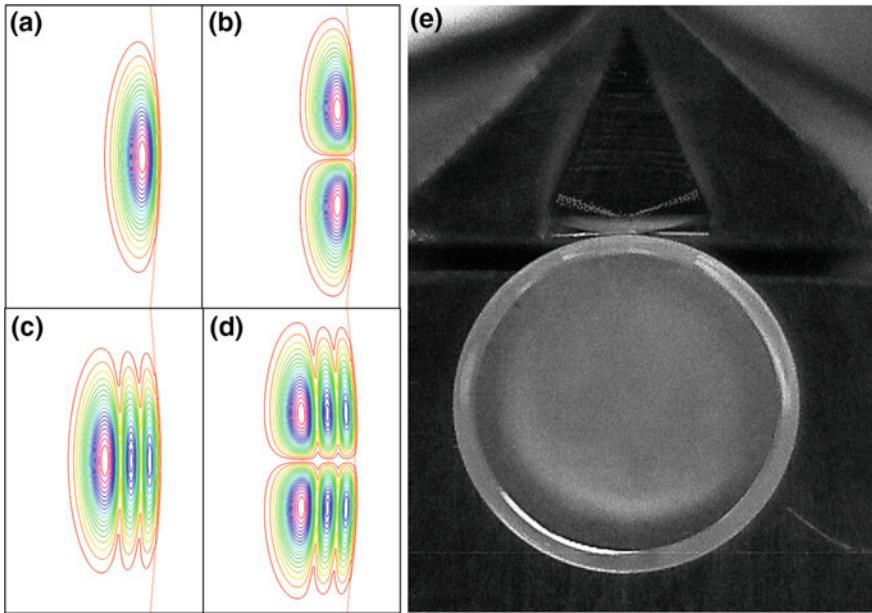


Fig. 3.8 Intensity distribution in the (r, θ) cross section of a WGM resonator for the fundamental mode $q = 1, p = 0$ (a), higher-order modes $q = 1, p = 1$ (b), $q = 3, p = 0$ (c), $q = 3, p = 1$ (d); and the top (r, φ) view of a resonator with the coupling prism (e). Optical beams, visible inside the prism because of fluorescence, are focused at the coupling region where the total internal reflection is locally frustrated

butions in the fundamental and three higher-order WGMs are shown in Fig. 3.8a–d. Coupling of WGM resonators to external optical beams is usually done via frustrated total internal reflection, which is achieved by placing a higher-index waveguide or prism in the evanescent field of the resonator, see Fig. 3.8e.

More detailed discussion of WGM resonators and their properties can be found in review papers [222–232]. Here we only make two comments regarding WGM resonators that are relevant to our topic. First, the quality factor Q of WGM resonators made from optically nonlinear crystals typically ranges from 10 to 100 millions. For a resonator with 1 mm circumference and 1 μm wavelength this translates to the finesse $\mathcal{F} = Q/m \sim 10^4\text{--}10^5$. Limited mainly by absorption of the material, high Q persists within its entire transparency window, which for a good optical crystal may well exceed an octave. Therefore the pump, signal and idler are all high-finesse modes, which increases the nonlinear optical conversion efficiency by a factor of $\mathcal{F}^3 \sim 10^{12}\text{--}10^{15}$ compared to the same millimeter-long crystal without a cavity. This is a very strong enhancement which allows to seriously discuss the perspectives of nonlinear and quantum optics with a few or even single photons, in particular implementing optical quantum logic gates [100].

The second note concerns the SPDC phase matching. While the formalism (3.10), (3.11), (3.12) still applies, the overlap integral (3.12) leads to selection rules that are much less restrictive than the usual phase matching (3.9). In fact the angular part of this integral yields the Clebsch-Gordan coefficients, reminding us that in spherical geometry the orbital momenta are conserved, rather than linear momenta. The radial part leads to no strict selection rules, but it favors such combinations when $q_p \approx q_s + q_i$ [233].

SPDC was observed in WGM resonators made from various optically nonlinear crystals and at various pump wavelengths both above [33, 234–239] and below [121, 124, 190, 240] the OPO threshold which for such resonators can be as low as several microwatts [235]. Two-mode squeezing above the threshold was reported by Fürst et al. [33]. The emitted signal and idler wavelengths can be tuned in a very wide range but at the same time with a great precision using a combination of temperature tuning, pump mode selection and evanescent field manipulation. Adjusting these parameters, Schunk et al. have been able to tune the signal wavelength to an atomic transition and observe fluorescence induced by single heralded photons [121]. In this experiment both cesium and rubidium D1 transitions were accessed using the same laser and the same resonator with the resonator temperature change by less than 2 °C.

Narrow linewidth of WGMs leads to a relatively sparse spectrum. Leveraging the selection rules, this can be used for engineering a single-mode parametric light source. Strictly single mode operation attested to by a Glauber correlation function measurement on the signal beam $g^{(2)}(0) = 2.01 \pm 0.07$ was demonstrated by Förtsch et al. with only minimal spectral filtering [124]. In this experiment the spectral width of the pulsed pump was transform-limited to approximately 20 MHz, exceeding the signal and idler spectral widths (both equal to the resonator linewidth) by more than a factor of two. Hence even a very careful measurement of the signal frequency would not allow to identify its idler twin photon among the others using relation (3.8), and true single-mode regime is achieved.

By the same argument, single-mode operation should not be expected with a CW pump having a linewidth smaller than that of a resonator mode. However an experiment using a sub-kHz wide CW laser pumping a WGM resonator with several MHz linewidth [190] showed surprisingly few (approximately three, where it should be thousands) SPDC modes, consistently with $g^{(2)}(0) \approx 1.3$. Note that in this experiment the “parasitic” SPDC into a wrong family of signal and idler WGMs has not been filtered out. Such filtering has improved $g^{(2)}(0)$ from 1.5 to 2 in the pulsed light experiment [124], see above. Therefore the extra modes observed in [190] are more likely to be associated with different mode families than with photons distinguishability within a single WGM.

The apparent paradox is resolved if we contemplate the fact that limitation of the observation time prevents us, even in principle, from performing a frequency measurement of the signal photon with the resolution required to localize the idler photon within a WGM linewidth. In this respect, gating a photon-detection measurement is equivalent to pulsing the pump. In both experiments [124, 190] the measurement time was defined by the resolution of the instrument recording the signal-idler

coincidences, 1 ns and 162 ps respectively, much too short for resolving the WGM linewidth.

Closing this section we would like to make two remarks regarding the cavity-assisted nonlinear optical processes. The first one is that squeezing can be attained not only in PDC but also in other such processes. For example, both the second harmonic [30, 241, 242] and fundamental pump wavelength [243, 244] in the frequency-doubling processes may be squeezed. But in such processes the amount of squeezing is inherently limited and most likely does not present significant practical interest. The second remark is that the parametric down-conversion near degeneracy may populate multiple pairs of quantum-correlated signal and idler modes [245, 246], leading to an optical comb. Such quantum-correlated optical combs may be used for creating multipartite entangled states, highly desired in many quantum information applications, e.g. in linear quantum computing. Finally, we would like to point out that WGM is not the only type of the optically nonlinear monolithic resonators based on total internal reflection. OPO based on square-shaped monolithic resonators has been recently implemented to generate 2.6 dB of vacuum squeezing [247].

3.2.3 Kerr Nonlinearity in Fibers and Resonators

A monochromatic wave propagating in Kerr media experiences self-phase modulation (SPM) that can be described by the Kerr Hamiltonian

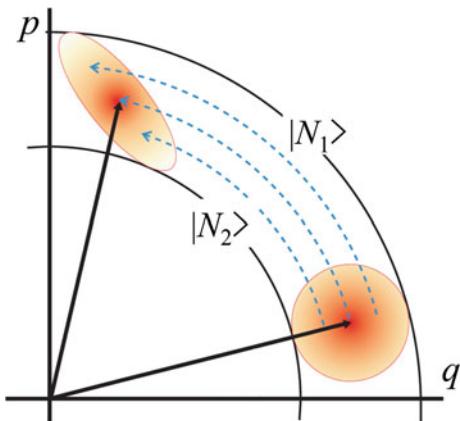
$$H = \hbar K a^\dagger a a^\dagger a \quad (3.14)$$

and by the associated time evolution operator [248]. If the nonlinear phase shift is small enough this interaction can be approximated by a dependence of the index of refraction n on the intensity I [249]:

$$n = n_0 + n_2 I. \quad (3.15)$$

Relation (3.15) is applicable to classical and quantum fluctuations of intensity. Expanding e.g. a coherent state in the photon-number basis we observe that the SPM advances a higher-number state $|N_1\rangle$ further in the phase space than a lower-number state $|N_2\rangle$. As a result, a characteristic shearing of the Wigner function occurs, as illustrated in Fig. 3.9, eventually leading to a crescent shape similar to Fig. 3.1f and indicating the number-phase squeezing [250]. The direction of shearing is opposite for materials with self-focusing ($n_2 > 0$) and self-defocusing ($n_2 < 0$). Note that SPM broadens the optical spectrum, leading to generation of frequency-shifted fields, but preserves the initial field energy. This process can be also described as degenerate four-wave mixing; in continuous-spectrum systems there is no clear boundary between these two processes.

Fig. 3.9 Illustration of an input coherent state squeezing via self phase modulation



Broad-band Kerr response in transparent dielectrics is much weaker than the resonant Kerr response in atoms, or the quadratic response in optical crystals. However, the Kerr nonlinearity in dielectrics has an important advantage: it is present also in amorphous materials such as fused silica, that can be shaped into long single-mode fibers that have very low loss. This advantage allowed Shelby et al. [251] to observe Kerr squeezing in fiber already in 1986, the same year as the first OPO squeezing was reported and a year after the first demonstration of squeezing in a sodium beam. They used 114 m of liquid helium cooled single-mode optical fiber pumped with CW 647 nm laser light. Reflecting the output light off a single-ended cavity they varied the phase between the pump (also serving as the local oscillator) and the squeezed sideband to observe 0.6 dB of squeezing. Liquid helium had to be used to suppress stimulated Brillouin oscillations and spontaneous guided acoustic-wave Brillouin scattering (GAWBS), the acousto-optic phenomena presenting the main obstacles to CW Kerr-squeezing in fibers.

These obstacles can be circumvented by using short pulses and high peak intensities. Because of different power dependence of the Kerr and Brillouin responses this effectively minimizes the latter. Bergman and Haus observed 5 dB of squeezing with 100-ps pulses propagating in a 50 m fiber loop Sagnac interferometer [252]. Alleviating the problem with GAWBS, short pulses bring about a difficulty of their own: GVD causes them to spread, losing the advantage of high peak power. This problem can be solved using optical solitons. Rosenbluh and Shelby have detected a modest (1.1 dB) squeezing of 200-fs soliton pulses propagating at room temperature in 5 m of optical fiber symmetric Sagnac interferometer [253]. Asymmetric Sagnac interferometers were later used to produce stronger amplitude squeezing of solitons: 3.9 dB (6.0 dB corrected for losses) with 126-fs pulses [254], and 5.7 dB (6.2 dB corrected for losses) with 182-fs pulses [255].

Sagnac loops are convenient because they naturally facilitate a homodyne measurement. However, detecting the photon-number squeezing in a direct measurement is also possible. This was accomplished in a unidirectionally pumped 1.5 km fiber,

yielding 2.3 dB (3.7 dB corrected for losses) squeezing of 2.3-ps soliton pulses [256]. In combination with the propagation length dependent spectral filtering, this technique has lead to even stronger (3.8 dB) squeezing of 130-fs pulses [257]. Squeezing bandwidth in this experiment is shown to be at least 2 GHz. Even higher bandwidths are theoretically possible. It is furthermore possible to generate mid-infrared time-locked patterns of squeezed vacuum with the amplitude fluctuations varying from below to above the shot noise limit, i.e. from squeezing to anti-squeezing, on the sub-cycle time scale [258]. Observation of this phenomenon is enabled by the sub-cycle electro-optic probing [259, 260].

The benefit of squeezing solitons does not come entirely for free: solitonic propagation requires specific input pulse shape and area, which makes the squeezing depend on the pulse energy [256]. But stabilizing the pulse energy is a much more tractable problem than suppressing GAWBS and managing GVD. And in addition, if the input energy is large enough for the given pulse parameters, the nonlinear dynamical evolution of the the pulse will lead to a soliton solution.

Polarization squeezing can be prepared from quadrature squeezing of two orthogonally polarized modes by projecting them onto a new polarization basis. Levandovsky et al. used for this purpose polarization-maintaining (PM) optical fibers in Sagnac configuration, producing about 1 dB of squeezing [261]. Better results were obtained with a unidirectionally pumped 13.3 m PM fiber [262], in which case 130-fs soliton pulses were squeezed to 5.1 dB. This result was later improved to 6.8 dB (10.4 dB corrected for losses), but Raman scattering was found to become a limiting factor at that level [263].

An interesting approach was taken by Margalit et al. [264], who used *off-diagonal* components of the $\chi^{(3)}$ tensor to cross-phase modulate orthogonal polarizations. In this case linearly polarized 1-nJ 150-fs pulses propagating unidirectionally in a non-PM fiber induced 3 dB of vacuum squeezing in the orthogonal polarization.

Invention of microstructured, hollow-core and photonic crystal fibers opened new opportunities in Kerr squeezing. In microstructured fibers, light is confined primarily in a thin solid core which concentrates the optical field in a smaller volume and increases the Kerr interaction strength. Furthermore, GVD in such fibers can be engineered by designing the structure around the core. Pumping a microstructured fiber near its zero GVD with 38-fs pulses, Hirosawa et al. [265] observed a spectrally broadened optical signal with up to 4.6 dB (10.3 dB corrected for losses) squeezing for some sidebands. Milanovic et al. [266] observed 3.9 dB of squeezing and a reduction of excess noise, i.e. an increase in purity, as compared to standard fiber squeezing experiments. Four wave mixing in microstructured fibers has been also used to create pulsed photon pairs at a rate rivaling the best SPDC sources [267, 268].

Another opportunity lies in combining the benefits of strong Kerr response of atomic transition with the field confinement and GVD engineering accessible in hollow-core optical fibers, in particular those with cross section resembling a traditional Japanese woven basket, which earned them a nickname Kagome fibers, see Fig. 3.10. In Kagome fibers, light propagates mainly inside the central hollow channel, which can be filled with a Kerr media of choice. GVD can still be tailored by designing the fiber microstructure surrounding the channel, but it can furthermore by

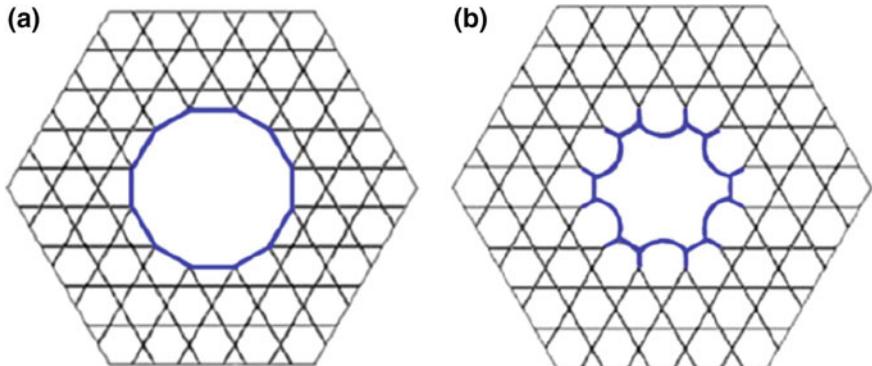


Fig. 3.10 Kagome fibers with a central channel designed for gas filling: a circular channel contour (**a**) and a hypocycloidal contour (**b**) designed to minimize the optical field diameter

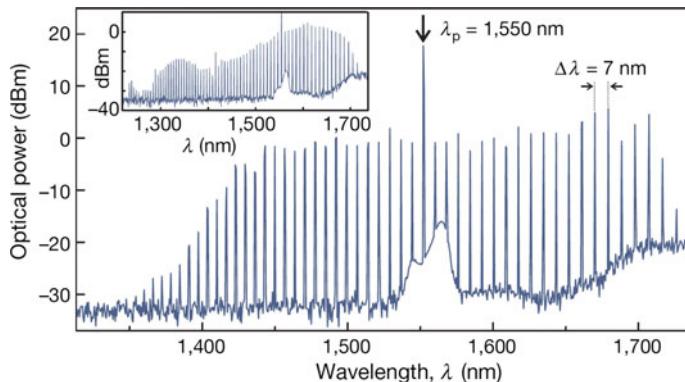


Fig. 3.11 A Kerr comb generated in a WGM resonator. Reprinted from [274]

dynamically fine-tuned by changing the gas pressure, literally inflating the Kagome fiber during the drawing process or even during the measurement [269]. At the same time, Brioullin and Raman processes in the fiber material are virtually avoided. First results have demonstrated squeezing in fibers filled with high pressure argon [270] and mercury vapour [271]. Filling Kagome fibers with alkali atom vapors has been proposed [272] and attempted, but has not yet led to success because of chemically aggressive properties of such vapors.

Extended interaction of strongly confined optical fields can be achieved not only in fibers, but also in resonators. In contrast to waveguides, resonators have discrete spectra consisting of nearly-equidistant modes. In this case the SPM, cross-phase modulation (XPM) and four-wave mixing processes are clearly distinct. All these processes play their roles in the formation of Kerr-combs in crystalline WGM resonators [273], such as shown in Fig. 3.11.

WGM combs have been extensively discussed recently, see e.g. [232, 275] and references therein. The aspect that is directly relevant to our discussion is the photon-number correlation between multiple pairs of sidebands placed symmetrically on both sides of the pump wavelength labeled λ_p in Fig. 3.11. This correlation arises from the degenerate four-wave mixing (or hyperparametric) process of annihilation of two pump photons and creation a photon pair in two symmetric modes. Below the oscillation threshold this process leads to the generation of entangled photon pairs. A number of experimental demonstrations of such pairs has emerged recently using on-chip fabricated silicon microring resonators [276–282]. The time-energy entanglement was proved by violating Bell’s inequality in [280, 281], and [282] has demonstrated time-energy and polarization hyper-entanglement, also confirmed by Bell’s inequality violation.

Above the threshold, hyperparametric conversion leads to two-mode squeezing in a multitude of mode pairs. Such squeezing was demonstrated in a microfabricated Si_3N_4 ring [283], which is not strictly speaking a WGM resonator, but is closely related. The free spectral range of this resonator $\Delta\lambda$ was large enough to allow selection of a single pair of squeezed modes by spectral filtering. These modes were found to be squeezed at the level of 1.7 dB (5 dB corrected for losses). Broadband quadrature squeezing in a similar resonator has been theoretically predicted [284].

Closing this section, we would like to mention that interaction of light with mechanical vibrations is not always harmful for preparation of non-classical light as in the case of GAWBS. It can be used to one’s advantage. Recently, it was shown that squeezed light can be created by coupling light with a mechanical oscillator. Here the radiation pressure quantum fluctuations induce the resonator motion which in turn imparts a phase shift to the laser light. Intensity-dependent phase shift leads to optical squeezing in close analogy to the Kerr effect. In this way squeezing of 1.7 dB was demonstrated in a bulk cavity setup containing a thin partially transparent mechanical membrane [285, 286].

3.2.4 Lasers and Other Feedback Systems

Laser light is commonly believed to be the best real-world approximation of a coherent state of an optical mode. However this is not always the case. The nonlinear response of a laser cavity can lead to sub-Poisson statistics of the emitted light, i.e. photon-number squeezing illustrated in Fig. 3.1f. To understand the physical mechanisms of intensity fluctuation suppression in lasers, consider an experiment with a vacuum tube filled with mercury vapor, carried out in 1985 [287]. In this experiment, a constant current flowing through the tube caused the fluorescence with the photon rate fluctuations below the shot noise. While the electrons emitted from the cathode have Poisson statistics, their flow through the vacuum tube is regulated by both the anode potential and the space charge of the electron flow. If the current increases, so does the negative space charge, which leads to the current fluctuation suppression. In other words, the space charge acts as a compressible buffer, smoothing out

these fluctuations below the classical limit, which is reflected in the emitted photons statistics. This is the same mechanism, which allowed Schottky and Spehnke [288] to observe a sub shot noise electron current in a vacuum tube in 1937.

A similar mechanism is present in semiconductor lasers operating in the constant-current (but not in the constant-voltage) regime, where the junction voltage provides a negative feedback regulating the current in the region of recombination [289]. This experiment was carried out using laser diodes at room temperature [290] and at 77 K [291]. In both cases approximately 1.7 dB amplitude squeezing (corrected for detectors efficiency) was detected in a very broad frequency range. Evidently, the squeezing measurement in these experiments was impeded by low collection efficiency. Improving this efficiency by “face-to-face” coupling of the laser diode and the photo diode, and cooling the assembly down to 66 K, the same group was able to boost the squeezing to 8.3 dB [292]. Considering the 89% quantum efficiency of the photodiode, this corresponds to 14 dB inferred squeezing. However, neither this nor other groups were later able to reproduce this large squeezing in a semiconductor laser, showing that there must be parameters not well understood and controlled in the initial experiment. Nevertheless their experiment initiated work in other groups, which eventually led to a better understanding.

Although the space charge model gives a qualitative understanding of the phenomenon, it does not capture many important details. In 1995, Marin et al. conceded that “the very mechanisms capable of explaining why some laser diodes and not others were generating sub-shot-noise light remained unclear” [293]. They came to the conclusion that one of these mechanisms is the cross-talk between the main mode and other weakly excited modes, which should lead to their anti-correlation, i.e. two-mode or even multipartite squeezing. Later, the same group developed a theoretical understanding by identifying two excess noise sources, the Petermann excess noise and the leakage current noise, to explain the limitations of the squeezing observed [294].

Another relevant factor is the optical injection into the laser cavity. The effects of an external laser injection at 10 K [295] and self-injection at room temperature [296] were studied in quantum-well lasers. Over 3 dB [295] and 1.8 dB [296] photon-number squeezing was observed. A weak squeezing in a free-running quantum-well laser was also observed at room temperature [297, 298].

The negative feedback suppressing the current (and hence the optical power) fluctuations does not necessarily have to be facilitated by the laser cavity. In Sect. 3.2.2 we already discussed an example of the electronic feedback derived from the signal measurement to control the idler photon statistics in PDC. A similar technique was applied to a semiconductor laser in 1986 by Yamamoto et al. [299]. Because the laser beam lacks a quantum-correlated twin, Yamamoto employed a XPM-based quantum nondemolition (QND) measurement to monitor the output laser power. The power fluctuations of the laser beam were imprinted onto the phase of a probe beam, recovered in a heterodyne measurement, and fed back to the laser current. As a result, the amplitude squeezing ranging from 5 dB at 16 MHz to 10 dB below 2 MHz was observed.

It might seem that a linear beam splitter could provide a simpler alternative to a QND measurement in preparation of non-classical light with the feedback technique. This approach indeed leads to a very interesting field dynamics known as *squashing* [300]. The term “squashing” pertains to the fields propagating inside the loop, and is fundamentally different from squeezing. The most remarkable property of the in-loop squashed optical field, theoretically shown by Shapiro et al. [301], is that such a field does not obey the usual commutation relations. Therefore it is not subject to Heisenberg uncertainty principle, and its photon-number uncertainty can be reduced below the classical limit without the phase noise penalty [300]. It is worth noting that not only a state of an optical mode, but also a motional state of a trapped ion can be squashed in a feedback loop [302].

In the context of nonclassical light applications, the possibility of generating optical fields not constrained by the Heisenberg uncertainty relations appears too good to be true. And indeed, it has been shown that out-coupling the squashed field from the loop destroys its remarkable properties [301]. In fact, it has been pointed out [300] that even fully characterizing these properties, which is only possible within the loop, is a highly nontrivial experimental problem that requires a QND measurement. Therefore using the electronic feedback systems for generating non-classical light has not attracted much of practical interest. Using feedforward, on the other hand, is quite common in commercial optical devices known as “noise eaters” that can suppress power fluctuations within the classical limit.

It would seem that diode lasers offer the most robust and easily scalable technology for generating non-classical (photon-number squeezed, or sub-Poissonian) light. They have also shown a promise in generating strongly squeezed states [292]. However the interest to this field apparently waned in the first decade of the 21st century. The reason for this skepticism could be that the discovery of the excess noise sources by Maurin et al. [294] made it clear that it is difficult to fabricate a laser that would *predictably* generate strong squeezing. If this is the case, a new advance in the field may be expected from improving the semiconductor technology.

3.3 Final Remarks

Non-classical light has played an important role in development of quantum theory, starting from the early tests of local realism performed with entangled photons in 1972 [137]. Following this pioneering experiment, many striking quantum phenomena have been discovered via non-classical optics research. Fluctuations of the optical field intensity have been suppressed below the shot-noise limit, which in classical notation requires negative probabilities. The concept of a *biphoton*, and later of a multipartite entangled state, was proven to be tangible. Thus physicists gained hands-on experience with a system that may consist of space-like separated parts and yet constitute a single physical entity. Experimental *quantum teleportation* has been made possible with such systems.

Not only fundamental, but also applied science and technology have a lot to benefit from non-classical light. Sub shot noise characteristics of the squeezed light directly points to one group of such applications: high resolution metrology. Optical phase in an interferometer, optical beam displacement, sub-wavelength image discerning and recording are just a few topics from this group. Information encoded in non-orthogonal single photon states or in any other non-orthogonal pure quantum states is protected from copying by fundamental laws of physics, which gives rise to another large group of applications concerned with information security. Furthermore, this information can be processed using mind-blowing quantum logic operations (such as e.g. a $\sqrt{\text{NOT}}$ gate) allowing, in perspective, to realize a quantum computer and the quantum internet.

But how is this wonderful non-classical light generated? The purpose of this chapter has been to provide a brief introductory tour over the most common sources of quantum light. The variety of physical systems capable of generating non-classical light is very broad. We encountered atomic beams, vapor cells, laser-cooled atomic clouds and even individual trapped atoms or ions; optical crystals and fibers; semiconductor nanoparticles and diode lasers.

With such a great variety of physical systems to discuss, we did not have an opportunity to provide much of detail regarding each system and its performance. Instead, we rely on references that are strategically placed so that an interested reader would be able to easily “zoom in” on any part of our review by downloading the appropriate publications.

Despite the great diversity of the quantum light sources, a few common properties can be summarized that are important for the majority of non-classical light applications. They are the following.

- *Optical nonlinearity.* This is a driving mechanism for generating non-classical light. Strongly nonlinear optical systems require less pump power and as a consequence are less noisy and more technologically acceptable. Resonant nonlinearity of natural or artificial atoms, and broad-band nonlinearity of laser gain media are two examples that may surpass other systems by far.
- *Optical loss.* When photons are randomly removed from the system, statistics of the remaining photons becomes more and more Poissonian. For many (but not all) quantum states this results in diminishing their non-classical characteristics, such as e.g. squeezing. For quantum states with zero displacement in phase space such as Fock-states, squeezed vacuum states, cat states (i.e. superpositions of coherent states) the statistical loss of a photon on average is enough to largely reduce the non-classical property [303, 304]. Good examples of low-loss systems are optical nonlinear crystals and fibers.
- *Mode structure.* While imaging applications require multiple transverse modes, the applications concerned with sensing and information processing may require strictly single-mode light. Bulk nonlinear crystals are natural sources for multi-mode light; on the other hand waveguides, fibers and optical cavities can be used to achieve single-mode operation.

- *Wavelength and bandwidth.* Using non-classical light in conjunction with “massive” qubits, as suggested by the quantum network paradigm, requires matching their central wavelengths and bandwidths. Therefore these source parameters either need to be precisely engineered, which is possible with atom-based sources, or tunable. Wavelength tuning is readily available in bulk crystals, but their emission is usually very broad-band. Fine-tuning of parametric light to atomic transitions in both central wavelength and bandwidth has been achieved with sub- or above-threshold OPOs.
- *Practical utility.* It is generally desirable to avoid cryogenic temperatures and other stringent environment requirements. Unfortunately, many of quantum dot, quantum well and trapped atoms sources of non-classical light fail this requirement. Therefore, progress may come via two different routes: (1) improving room temperature systems, or (2) developing compact sources and low cost cryogenic fridges.

In general, we see that progress in quantum optics comes from developing: (1) light sources, (2) light confinement strategies, (3) materials with strong optically nonlinear response. For a particular goal, one achieves the best result if one optimizes the combination of items from these three categories. We have already discussed one such example, a hollow-core fiber filled with atom media, and it appears plausible that more such examples may emerge in the future.

Acknowledgements We thank Drs. M. Raymer and M. Gurioli for valuable comments. D. V. S. would like to thank the Alexander von Humboldt Foundation for sponsoring his collaboration with the Max Plank Institute for the physics of light in Erlangen.

Appendix: Is Coherent Light Quantum?

Let us consider the following series of thought experiments. The toolbox we need contains a source of laser light, a beam splitter, two time resolving detectors of high bandwidth, and electronic equipment to analyze the detector signals. In the first experiment (1) measuring the intensity correlations after splitting the laser light with the beam splitter yields a $g^{(2)}(\tau)$ which is independent of time τ . This can be described by a classical model, namely classical light fields without fluctuations—fine. Now the second experiment (2) is to measure the intensity of the laser light as a function of time. The result is a fluctuating detector signal (corresponding to the Poisson statistics of the photons in a quantum language). A classical model can also describe this. This time it is a model in which the classical electric fields fluctuate—this is also fine, but note that the models required are not compatible.

You may not be satisfied and argue that the fluctuation observed in experiment (2) may well come from the detectors themselves contributing noise. This would average out in experiment (1) because the noises introduced by the two detectors are of course not correlated. But suppose the lab next door happens to have amplitude squeezed light, with intensity fluctuations suppressed by 15 dB below the shot noise.

Measuring the squeezed light intensity noise you convince yourself easily that the detector does not introduce enough noise to explain experiment (2). Note that this test should convince you even if you have no clue what the squeezed light is.

But you do not want to give up so easily and you say “what if a classically noisy light field enters the second input port, uncorrelated with the laser light but likewise modeled by classical stochastic fluctuations?”. And you are right, this more involved classical model would explain both experiments (1) and (2)—yet there is (3) a third experiment we can do. We can check the intensity of the light arriving at this second input port of the beam splitter and no matter how sensitive the intensity measuring detectors are they will detect no signal. But this is not compatible with a classical model: classical fluctuations always lead to measurable intensity noise.

We conclude by noting that obviously coherent states are non-classical because there is no single classical stochastic model which describes all possible experiments with laser light. But as we have seen it is tedious to go through these arguments, and no simple measure of non-classicality was found so far qualifying a coherent state as non-classical. Nevertheless, the non-classical nature of a coherent state is used in some quantum protocols.

It is interesting to note that there is a much different scenario in which experiments with coherent states cannot be described classically without field quantization, i.e. with semi classical theory. Coherent states lead e.g. to a revival of Rabi oscillations in their interaction with an atom in the Jaynes Cummings model. This effect can only be properly described when properly accounting for the quantization of the electromagnetic field [305, 306]. Thus the hypothesis is that for any pure quantum state it is always possible to find experimental scenarios, which can only be properly described using field quantization. Let us furthermore note that also thermal states, i.e. mixed quantum states, can still be somewhat nonclassical in nature if the classical excess noise is not too much larger than the underlying quantum uncertainty.

References

1. R. Loudon, *The Quantum Theory of Light* (Oxford University Press, 2000)
2. D.N. Klyshko, Basic quantum mechanical concepts from the operational viewpoint. Phys.-Uspekhi **41**, 885–922 (1998)
3. D.F. Walls, Evidence for the quantum nature of light. Nature **280**, 451–454 (1979)
4. H. Paul, Photon antibunching. Rev. Mod. Phys. **54**, 1061–1102 (1982)
5. G. Leuchs, Photon statistics, antibunching and squeezed states, in *Frontiers of Nonequilibrium Statistical Physics*, ed. by G.T. Moore, M.O. Scully (Springer, US, 1986)
6. D.N. Klyshko, The nonclassical light. Phys.-Uspekhi **39**, 573–596 (1996)
7. H.J. Kimble, M. Dagenais, L. Mandel, Photon antibunching in resonance fluorescence. Phys. Rev. Lett. **39** (1997)
8. D.N. Klyshko, Quantum optics: quantum, classical, and metaphysical aspects. Phys.-Uspekhi **37**, 1097–1123 (1994)
9. J.S. Bell, On the einstein podolsky Rosen paradox. Physics **1**, 195–200 (1964)
10. J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, Proposed experiment to test local hidden-variables theories. Phys. Rev. Lett. **23**, 880–884 (1969)

11. J.F. Clauser, M.A. Horne, Experimental consequences of objective local theories. *Phys. Rev. D* **10**, 526–535 (1974)
12. J.F. Clauser, A. Shimony, Bell’s theorem: experimental tests and implications. *Rep. Prog. Phys.* **41**, 1881–1927 (1978)
13. N.J. Cerf, C. Adami, Negative entropy and information in quantum mechanics. *Phys. Rev. Lett.* **79**, 5194–5197 (1997)
14. M.A. Nielsen, I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010)
15. W.P. Schleich, *Quantum Optics in Phase Space* (Wiley-VCH Verlag Berlin GmbH, Berlin, 2001)
16. M. Hillery, R.F. O’Connell, M.O. Scully, E.P. Wigner, Distribution functions in physics: fundamentals. *Phys. Rep.* **106**, 121–167 (1984)
17. M. Hillery, Total noise and nonclassical states. *Phys. Rev. A* **39**, 2994–3002 (1989)
18. C.T. Lee, Higher-order criteria for nonclassical effects in photon statistics. *Phys. Rev. A* **41**, 1721–1723 (1990)
19. D.N. Klyshko, Observable signs of nonclassical light. *Phys. Lett. A* **213**, 7 (1996)
20. B. Yurke, D. Stoler, Generating quantum mechanical superpositions of macroscopically distinguishable states via amplitude dispersion. *Phys. Rev. Lett.* **57**, 13–16 (1986)
21. B. Vlastakis, G. Kirchmair, Z. Leghtas, S.E. Nigg, L. Frunzio, S.M. Girvin, M. Mirrahimi, M.H. Devoret, R.J. Schoelkopf, Deterministically encoding quantum information using 100-photon schrödinger cat states. *Science* **342**, 607–610 (2013)
22. P.G. Kwiat, H. Weinfurter, Embedded bell-state analysis. *Phys. Rev. A* **58**, R2623–R2626 (1998)
23. J. Dowling, Quantum optical metrology—the lowdown on hing-noon states. *Contemp. Phys.* **49**, 125–143 (2008)
24. I. Afek, O. Ambar, Y. Silberberg, High-noon states by mixing quantum and classical light. *Science* **328**, 879–881 (2010)
25. Y. Zhang, T. Furuta, R. Okubo, K. Takahashi, T. Hirano, Experimental generation of broadband quadrature entanglement using laser pulses. *Phys. Rev. A* **76**, 012314 (2007)
26. K.-I. Yoshino, T. Aoki, A. Furusawa, Generation of continuous-wave broadband entangled beams using periodically poled lithium niobate waveguides. *Appl. Phys. Lett.* **90**, 041111 (2007)
27. U.L. Andersen, G. Leuchs, C. Silberhorn, Continuous-variable quantum information processing. *Las. Phot. Rev.* **4**, 337–354 (2010)
28. W.P. Bowen, N. Treps, B.C. Buchler, R. Schnabel, T.C. Ralph, H.-A. Bachor, T. Symul, P.K. Lam, Experimental investigation of continuous-variable quantum teleportation. *Phys. Rev. A* **67**, 032302 (2003)
29. M.V. Chekhova, G. Leuchs, M. Żukowski, Bright squeezed vacuum: entanglement of macroscopic light beams. *Opt. Commun.* **337**, 27–43 (2015)
30. M.J. Collett, D.F. Walls, Squeezing spectra for nonlinear optical systems. *Phys. Rev. A* **32**, 2887–2892 (1985)
31. M.D. Reid, P.D. Drummond, Quantum correlations of phase in nondegenerate parametric oscillation. *Phys. Rev. Lett.* **60**, 2731–2733 (1988)
32. C. Fabre, E. Giacobino, A. Heidmann, S. Reynaud, Noise characteristics of a non-degenerate optical parametric oscillator—application to quantum noise reduction. *J. Phys.* **50**, 1209–1225 (1989)
33. J.U. Fürst, D.V. Strekalov, D. Elser, A. Aiello, U.L. Andersen, C. Marquardt, G. Leuchs, Quantum light from a whispering-gallery-mode disk resonator. *Phys. Rev. Lett.* **106**, 113901 (2011)
34. B. Yurke, S.L. McCall, J.R. Klauder, $Su(2)$ and $su(1,1)$ interferometers. *Phys. Rev. A* **33**, 4033–4054 (1986)
35. R.A. Campos, B.E.A. Saleh, M.C. Teich, Quantum-mechanical lossless beam splitter: $Su(2)$ symmetry and photon statistics. *Phys. Rev. A* **40**, 1371–1384 (1989)

36. K.Y. Spasibko, F. Töppel, T.S. Iskhakov, M. Stobiska, M.V. Chekhova, G. Leuchs, Interference of macroscopic beams on a beam splitter: phase uncertainty converted into photon-number uncertainty. *New J. Phys.* **16**, 013025 (2014)
37. D.M. Greenberger, M.A. Horne, A. Shimony, A. Zeilinger, Bell's theorem without inequalities. *Am. J. Phys.* **58**, 1131–1143 (1990)
38. J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, A. Zeilinger, Experimental test of quantum nonlocality in three-photon greenberger-horne-zeilinger entanglement. *Nature* **403**, 515–519 (2000)
39. W. Dur, G. Vidal, J.I. Cirac, Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A* **62**, 062314 (2000)
40. M. Eibl, N. Kiesel, M. Bourennane, C. Kurtsiefer, H. Weinfurter, Experimental realization of a three-qubit entangled w state. *Phys. Rev. Lett.* **92**, 077901 (2004)
41. J. Wen, S. Du, M. Xiao, Improving spatial resolution in quantum imaging beyond the rayleigh diffraction limit using multiphoton w entangled states. *Phys. Lett. A* **374**, 3908–3911 (2010)
42. H.J. Briegel, R. Raussendorf, Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.* **86**, 910–913 (2001)
43. M. Hein, J. Eisert, H.J. Briegel, Multipartite entanglement in graph states. *Phys. Rev. A* **69**, 062311 (2004)
44. J.A. Smolin, Four-party unlockable bound entangled state. *Phys. Rev. A* **63**, 032306 (2001)
45. H. Bechmann-Pasquinucci, W. Tittel, Quantum cryptography using larger alphabets. *Phys. Rev. A* **61**, 062308 (2000)
46. S.P. Walborn, D.S. Lemos, M.P. Almeida, P.H.S. Ribeiro, Quantum key distribution with higher-order alphabets using spatially encoded qudits. *Phys. Rev. Lett.* **96**, 090501 (2006)
47. P.B. Dixon, G.A. Howland, J. Schneeloch, J.C. Howell, Quantum mutual information capacity for high-dimensional entangled states. *Phys. Rev. Lett.* **108**, 143603 (2012)
48. W. Wasilewski, A.I. Lvovsky, K. Banaszek, C. Radzewicz, Pulsed squeezed light: simultaneous squeezing of multiple modes. *Phys. Rev. A* **73**, 063819 (2006)
49. D. Collins, N. Gisin, N. Linden, S. Massar, S. Popescu, Bell inequalities for arbitrarily high-dimensional systems. *Phys. Rev. Lett.* **88**, 040404 (2002)
50. H.-P. Lo, C.-M. Li, A. Yabushita, Y.-N. Chen, C.-W. Luo, T. Kobayashi, Experimental violation of bell inequalities for multi-dimensional systems. *Sci. Rep.* **6**, 22088 (2016)
51. A.C. Dada, J. Leach, G.S. Buller, M.J. Padgett, E. Andersson, Experimental high-dimensional two-photon entanglement and violations of generalized bell inequalities. *Nat. Phys.* **7**, 677–680 (2011)
52. A. Mair, A. Vaziri, G. Weihs, A. Zeilinger, Entanglement of the orbital angular momentum states of photons. *Nature* **412**, 313–316 (2001)
53. M. Krenn, M. Huber, R. Fickler, R. Lapkiewicz, S. Ramelow, A. Zeilinger, Generation and confirmation of a (100×100) -dimensional entangled quantum system. *PNAS* **111**, 6243–6247 (2014)
54. B.C. Hiesmayr, M.J.A. de Dood, W. Löffler, Observation of four-photon orbital angular momentum entanglement. *Phys. Rev. Lett.* **116**, 073601 (2016)
55. M.W. Mitchell, F.A. Beduini, Extreme spin squeezing for photons. *New J. Phys.* **16**, 073027 (2014)
56. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Quantum entanglement. *Rev. Mod. Phys.* **81**, 865–942 (2009)
57. G. Vidal, R.F. Werner, Computable measure of entanglement. *Phys. Rev. A* **65**, 032314 (2002)
58. W.K. Wootters, Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248 (1998)
59. R. Hildebrand, Concurrence revisited. *J. Math. Phys.* **48**, 102108–102108 (2007)
60. B.Y. Zeldovich, D.N. Klyshko, Statistics of field in parametric luminescence. *Sov. Phys. JETP Lett.* **9**, 40–44 (1969)
61. D.C. Burnham, D.L. Weinberg, Observation of simultaneity in parametric production of optical photon pairs. *Phys. Rev. Lett.* **25**, 84–87 (1970)

62. D.N. Klyshko, Use of two-photon light for absolute calibration of photoelectric detectors. *Quantum Electron.* **7**, 1932–1940 (1980)
63. S.V. Polyakov, A.L. Migdall, High accuracy verification of a correlated-photon-based method for determining photon-counting detection efficiency. *Opt. Express* **15**, 1390–1407 (2007)
64. M. Ware, A. Migdall, J. Bienfang, S. Polyakov, Calibrating photon-counting detectors to high accuracy: background and deadtime issues. *J. Mod. Opt.* **54**, 361–372 (2007)
65. A. Czitrovsky, A. Sergienko, P. Jani, A. Nagy, Measurement of quantum efficiency using correlated photon pairs and a single-detector technique. *Metrologia* **37**, 617–620 (2000)
66. M.V. Lebedev, A.A. Shchekin, O.V. Misochko, Two-electron pulses of a photomultiplier and two-photon photoeffect. *Quantum Electron.* **38**, 710–723 (2008)
67. G. Brida, M. Genovese, I. Ruo-Berchera, M. Chekhova, A. Penin, Possibility of absolute calibration of analog detectors by using parametric downconversion: a systematic study. *JOSA B* **23**, 2185–2193 (2006)
68. H. Vahlbruch, M. Mehmet, K. Danzmann, R. Schnabel, Detection of 15 dB squeezed states of light and their application for the absolute calibration of photoelectric quantum efficiency. *Phys. Rev. Lett.* **117**, 110801 (2016)
69. G. Brida, I.P. Degiovanni, M. Genovese, M.L. Rastello, I. Ruo-Berchera, Detection of multi-mode spatial correlation in PDC and application to the absolute calibration of a CCD camera. *Opt. Express* **18**, 20572–20584 (2010)
70. D.N. Klyshko, *Photons and Nonlinear Optics* (Taylor and Francis, New York, NY USA, 1988)
71. D.N. Klyshko, A.N. Penin, The prospects of quantum photometry. *Sov. Phys. Uspekhi* **30**, 716–723 (1987)
72. M. Xiao, L.-A. Wu, H.J. Kimble, Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.* **59**, 278–281 (1987)
73. P. Grangier, R.E. Slusher, B. Yurke, A. LaPorta, Squeezed-light-enhanced polarization interferometer. *Phys. Rev. Lett.* **59**, 2153–2156 (1987)
74. T.L.S. Collaboration, A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nat. Phys.* **7**, 962–965 (2011)
75. E.S. Polzik, J. Carri, H.J. Kimble, Spectroscopy with squeezed light. *Phys. Rev. Lett.* **68**, 3020–3023 (1992)
76. P.H.S. Ribeiro, C. Schwob, A. Maitre, C. Fabre, Sub-shot-noise high-sensitivity spectroscopy with optical parametric oscillator twin beams. *Opt. Lett.* **22**, 1893–1895 (1997)
77. M.A. Taylor, J. Janousek, V. Daria, J. Knittel, B. Hage, H.-A. Bachor, W.P. Bowen, Biological measurement beyond the quantum limit. *Nat. Phot.* **7**, 229–233 (2013)
78. J. Gea-Banacloche, Two-photon absorption of nonclassical light. *Phys. Rev. Lett.* **62**, 1603–1606 (1989)
79. J. Javanainen, P.L. Gould, Linear intensity dependence of a two-photon transition rate. *Phys. Rev. A* **41**, 5088–5091 (1990)
80. B. Dayan, Theory of two-photon interactions with broadband down-converted light and entangled photons. *Phys. Rev. A* **76**, 043813 (2007)
81. N.P. Georgiades, E.S. Polzik, K. Edamatsu, H.J. Kimble, A.S. Parkins, Nonclassical excitation for atoms in a squeezed vacuum. *Phys. Rev. Lett.* **75**, 3426–3429 (1995)
82. B. Dayan, A. Pe'er, A.A. Friesem, Y. Silberberg, Two photon absorption and coherent control with broadband down-converted light. *Phys. Rev. Lett.* **93**, 023005 (2004)
83. F. Boitier, A. Godard, E. Rosencher, C. Fabre, Measuring photon bunching at ultrashort timescale by two-photon absorption in semiconductors. *Nat. Phys.* **5**, 267–270 (2009)
84. D.Y. Korystov, S.P. Kulik, A.N. Penin, Rozhdestvenski hooks in two-photon parametric light scattering. *JETP Lett.* **73**, 214–218 (2001)
85. A.N. Boto, P. Kok, D.S. Abrams, S.L. Braunstein, C.P. Williams, J.P. Dowling, Quantum interferometric optical lithography: exploiting entanglement to beat the diffraction limit. *Phys. Rev. Lett.* **85**, 2733–2736 (2000)
86. A. Pe'er, B. Dayan, M. Vučelja, Y. Silberberg, A.A. Friesem, Quantum lithography by coherent control of classical light pulses. *Opt. Express* **12**, 6600–6605 (2004)

87. E.M. Nagasako, S.J. Bentley, R.W. Boyd, G.S. Agarwal, Nonclassical two-photon interferometry and lithography with high-gain parametric amplifiers. *Phys. Rev. A* **64**, 043802 (2001)
88. B. Dayan, A. Pe'er, A.A. Friesem, Y. Silberberg, Nonlinear interactions with an ultrahigh flux of broadband entangled photons. *Phys. Rev. Lett.* **94**, 043602 (2005)
89. T.B. Pittman, Y.H. Shih, D.V. Strekalov, A.V. Sergienko, Optical imaging by means of two-photon quantum entanglement. *Phys. Rev. A* **52**, R3429–R3432 (1995)
90. M.B. Nasr, B.E.A. Saleh, A.V. Sergienko, M.C. Teich, Demonstration of dispersion-canceled quantum-optical coherence tomography. *Phys. Rev. Lett.* **91**, 083601 (2003)
91. N. Treps, N. Grosse, W.P. Bowen, C. Fabre, H.-A. Bachor, P.K. Lam, A quantum laser pointer. *Science* **301**, 940–943 (2003)
92. R.P. Feynman, Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982)
93. P.W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comp.* **26**, 1484–1509 (1997)
94. J.F. Clauser, J.P. Dowling, Factoring integers with young's n-slit interferometer. *Phys. Rev. A* **53**, 4587–4590 (1996)
95. J.D. Franson, B.C. Jacobs, T.B. Pittman, Quantum computing using single photons and the zeno effect. *Phys. Rev. A* **70**, 062302 (2004)
96. J.D. Franson, T.B. Pittman, B.C. Jacobs, Zeno logic gates using microcavities. *JOSA B* **24**, 209–213 (2007)
97. B.D. Clader, S.M. Hendrickson, R.M. Camacho, B.C. Jacobs, All-optical microdisk switch using EIT. *Opt. Express* **21**, 6169–6179 (2013)
98. Y.-P. Huang, J.B. Altepeter, P. Kumar, Interaction-free all-optical switching via the quantum zeno effect. *Phys. Rev. A* **82**, 063826 (2010)
99. Y.-P. Huang, P. Kumar, Interaction-free all-optical switching in chi⁽²⁾ microdisks for quantum applications. *Opt. Lett.* **35**, 2376–2378 (2010)
100. Y.-Z. Sun, Y.-P. Huang, P. Kumar, Photonic nonlinearities via quantum zeno blockade. *Phys. Rev. Lett.* **110**, 223901 (2013)
101. S.M. Hendrickson, C.N. Weiler, R.M. Camacho, P.T. Rakich, A.I. Young, M.J. Shaw, T.B. Pittman, J.D. Franson, B.C. Jacobs, All-optical-switching demonstration using two-photon absorption and the zeno effect. *Phys. Rev. A* **87**, 23808 (2013)
102. D.V. Strekalov, A.S. Kowlgy, Y.-P. Huang, P. Kumar, Progress towards interaction-free all-optical devices. *Phys. Rev. A* **89**, 063820 (2014)
103. H.J. Kimble, The quantum internet. *Nature* **453**, 1023–1030 (2008)
104. T. Aoki, A.S. Parkins, D.J. Alton, C.A. Regal, B. Dayan, E. Ostby, K.J. Vahala, H.J. Kimble, Efficient routing of single photons by one atom and a microtoroidal cavity. *Phys. Rev. Lett.* **102**, 083601 (2009)
105. H.P. Specht, C. Nölleke, A. Reiserer, M. Uphoff, E. Figueroa, S. Ritter, G. Rempe, A single-atom quantum memory. *Nature* **473**, 190–193 (2011)
106. A. Ourjoumtsev, A. Kubanek, M. Koch, C. Sames, P.W.H. Pinkse, G. Rempe, K. Murr, Observation of squeezed light from one atom excited with two photons. *Nature* **474**, 623–626 (2011)
107. W. Chen, K.M. Beck, R. Bücker, M. Gullans, M.D. Lukin, H. Tanji-Suzuki, V. Vuletić, All-optical switch and transistor gated by one stored photon. *Science* **341**, 768–770 (2013)
108. S. Baur, D. Tiarks, G. Rempe, S. Dürr, Single-photon switch based on rydberg blockade. *Phys. Rev. Lett.* **112**, 073901 (2014)
109. X. Shomroni, S. Rosenblum, Y. Lovsky, O. Bechler, G. Guendelman, B. Dayan, All-optical routing of single photons by a one-atom switch controlled by a single photon. *Science* **345**, 903–906 (2014)
110. T.G. Tiecke, J.D. Thompson, N.P. de Leon, L.R. Liu, V. Vuletić, M.D. Lukin, Nanophotonic quantum phase switch with a single atom. *Nature* **508**, 241–244 (2014)
111. S. Rosenblum, O. Bechler, I. Shomroni, Y. Lovsky, G. Guendelman, B. Dayan, Extraction of a single photon from an optical pulse. *Nat. Phot.* **10**, 19–22 (2016)
112. P. Michler, A. Kiraz, C. Becher, W.V. Schoenfeld, P.M. Petroff, L. Zhang, E. Hu, A. Imamoglu, A quantum dot single-photon turnstile device. *Science* **290**, 2282–2285 (2000)

113. P.-B. Li, S.-Y. Gao, F.-L. Li, Quantum-information transfer with nitrogen-vacancy centers coupled to a whispering-gallery microresonator. *Phys. Rev. A* **83**, 054306 (2011)
114. Q. Chen, W.L. Yang, M. Feng, Quantum gate operations in decoherence-free fashion with separate nitrogen-vacancy centers coupled to a whispering-gallery mode resonator. *Eur. Phys. J. D* **66**, 238 (2012)
115. J. Volz, M. Weber, D. Schlenk, W. Rosenfeld, J. Vrana, K. Saucke, C. Kurtsiefer, H. Weinfurter, Observation of entanglement of a single photon with a trapped atom. *Phys. Rev. Lett.* **96**, 030404 (2006)
116. J. Beugnon, M.P.A. Jones, J. Dingjan, B. Darquié, G. Messin, A. Browaeys, P. Grangier, Quantum interference between two single photons emitted by independently trapped atoms. *Nature* **440**, 779–782 (2006)
117. P. Maunz, D.L. Moehring, S. Olmschenk, K.C. Younge, D.N. Matsukevich, C. Monroe, Quantum interference of photon pairs from two remote trapped atomic ions. *Nat. Phys.* **3**, 538–541 (2007)
118. V. Leong, S. Kosen, B. Srivathsan, G.K. Gulati, A. Cerè, C. Kurtsiefer, Hong-ou-mandel interference between triggered and heralded single photons from separate atomic systems. *Phys. Rev. A* **91**, 063829 (2015)
119. X.-H. Bao, Y. Qian, J. Yang, H. Zhang, Z.-B. Chen, T. Yang, J.-W. Pan, Generation of narrow-band polarization-entangled photon pairs for atomic quantum memories. *Phys. Rev. Lett.* **101**, 190501 (2008)
120. J. Fekete, D. Rieländer, M. Cristiani, H. de Riedmatten, Ultranarrow-band photon-pair source compatible with solid state quantum memories and telecommunication networks. *Phys. Rev. Lett.* **110** (2013)
121. G. Schunk, U. Vogl, D.V. Strekalov, M. Förtsch, F. Sedlmeir, H.G.L. Schwefel, M. Göbel, S. Christiansen, G. Leuchs, C. Marquardt, Interfacing transitions of different alkali atoms and telecom bands using one narrowband photon pair source. *Optica* **2**, 773–778 (2015)
122. A. Lenhard, M. Bock, C. Becher, S. Kucera, J. Brito, P. Eich, P. Müller, J. Eschner, Telecom-heralded single-photon absorption by a single atom. *Phys. Rev. A* **92**, 063827 (2015)
123. G. Schunk, U. Vogl, F. Sedlmeir, D.V. Strekalov, A. Otterpohl, V. Averchenko, H.G.L. Schwefel, G. Leuchs, C. Marquardt, Frequency tuning of single photons from a whispering-gallery mode resonator to MHz-wide transitions. *J. Mod. Opt.* **63**, 2058–2073 (2016)
124. M. Förtsch, G. Schunk, J.U. Fürst, D. Strekalov, T. Gerrits, M.J. Stevens, F. Sedlmeir, H.G.L. Schwefel, S.W. Nam, G. Leuchs, C. Marquardt, Highly efficient generation of single-mode photon pairs from a crystalline whispering-gallery-mode resonator source. *Phys. Rev. A* **91**, 023812 (2015)
125. K.-H. Luo, H. Herrmann, S. Krapick, B. Brecht, R. Ricken, V. Quiring, H. Suche, W. Sohler, C. Silberhorn, Direct generation of genuine single-longitudinal-mode narrowband photon pairs. *New J. Phys.* **17**, 073039 (2015)
126. E. Knill, R. Laflamme, G.J. Milburn, A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001)
127. Y.I. Bogdanov, M.V. Chekhova, L.A. Krivitsky, S.P. Kulik, A.N. Penin, A.A. Zhukov, L.C. Kwek, C.H. Oh, M.K. Tey, Statistical reconstruction of qutrits. *Phys. Rev. A* **70**, 042303 (2004)
128. B.P. Lanyon, T.J. Weinhold, N.K. Langford, J.L. O'Brien, K.J. Resch, A. Gilchrist, A.G. White, Manipulating biphotonic qutrits. *Phys. Rev. Lett.* **100**, 060504 (2008)
129. Y.I. Bogdanov, E.V. Moreva, G.A. Maslennikov, R.F. Galeev, S.S. Straupe, S.P. Kulik, Polarization states of four-dimensional systems based on biphotons. *Phys. Rev. A* **73**, 063810 (2006)
130. M.-X. Luo, Y. Deng, H.-R. Li, S.-Y. Ma, Photonic ququart logic assisted by the cavity-qed system. *Sci. Rep.* **5**, 13255 (2015)
131. W.K. Wootters, W.H. Zurek, A single quantum cannot be cloned. *Nature* **299**, 802–803 (1982)
132. C.H. Bennett, G. Brassard, Quantum cryptography: public key distribution and coin tossing, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol. 175 (1984), p. 8

133. S.L. Braunstein, P. van Loock, Quantum information with continuous variables. *Rev. Mod. Phys.* **77**, 513–577 (2005)
134. C. Crèpeau, J. Kilian, Achieving oblivious transfer using weakened security assumptions, in *29th Annual Symposium on Foundations of Computer Science* (1988), pp. 42–52
135. T. Lunghi, J. Kaniewski, F. Bussières, R. Houlmann, M. Tomamichel, A. Kent, N. Gisin, S. Wehner, H. Zbinden, Experimental bit commitment based on quantum communication and special relativity. *Phys. Rev. Lett.* **111**, 180504 (2013)
136. C. Croal, C. Peuntinger, B. Heim, I. Khan, C. Marquardt, G. Leuchs, P. Wallden, E. Andersson, N. Korolkova, Free-space quantum signatures using heterodyne measurements. *Phys. Rev. Lett.* **117**, 100503 (2016)
137. S.J. Freedman, J.F. Clauser, Experimental test of local hidden-variable theories. *Phys. Rev. Lett.* **28**, 938–941 (1972)
138. T. Brannan, Z. Qin, A. MacRae, A.I. Lvovsky, Generation and tomography of arbitrary optical qubits using transient collective atomic excitations. *Opt. Lett.* **39**, 5447–5450 (2014)
139. R.E. Slusher, L.W. Hollberg, B. Yurke, J.C. Mertz, J.F. Valleys, Observation of squeezed states generated by four-wave mixing in an optical cavity. *Phys. Rev. Lett.* **55**, 2409–2412 (1985)
140. A. Lambrecht, T. Coudreau, A.M. Steinberg, E. Giacobino, Squeezing with cold atoms. *Europhys. Lett.* **36**, 93–98 (1996)
141. C.F. McCormick, V. Boyer, E. Arimondo, P.D. Lett, Strong relative intensity squeezing by four-wave mixing in rubidium vapor. *Opt. Lett.* **32**, 178–180 (2007)
142. N. Corzo, A.M. Marino, K.M. Jones, P.D. Lett, Multi-spatial-mode single-beam quadrature squeezed states of light from four-wave mixing in hot rubidium vapor. *Opt. Express* **19**, 21358–21369 (2011)
143. V. Boyer, A.M. Marino, P.D. Lett, Generation of spatially broadband twin beams for quantum imaging. *Phys. Rev. Lett.* **100**, 143601 (2008)
144. V. Balić, D.A. Braje, P. Kolchin, G.Y. Yin, S.E. Harris, Generation of paired photons with controllable waveforms. *Phys. Rev. Lett.* **94**, 183601 (2005)
145. C.W. Chou, S.V. Polyakov, A. Kuzmich, H.J. Kimble, Single-photon generation from stored excitation in an atomic ensemble. *Phys. Rev. Lett.* **92**, 213601 (2004)
146. S.V. Polyakov, C.W. Chou, D. Felinto, H.J. Kimble, Temporal dynamics of photon pairs generated by an atomic ensemble. *Phys. Rev. Lett.* **93**, 263601 (2004)
147. M.D. Eisaman, L. Childress, A. André, F. Massou, A.S. Zibrov, M.D. Lukin, Shaping quantum pulses of light via coherent atomic memory. *Phys. Rev. Lett.* **93**, 233602 (2004)
148. J.F. Chen, S. Zhang, H. Yan, M.M.T. Loy, G.K.L. Wong, S. Du, Shaping biphoton temporal waveforms with modulated classical fields. *Phys. Rev. Lett.* **104**, 183604 (2010)
149. A.B. Matsko, I. Novikova, G.R. Welch, D. Budker, D.F. Kimball, S.M. Rochester, Vacuum squeezing in atomic media via self-rotation. *Phys. Rev. A* **66**, 043815 (2002)
150. S. Barreiro, P. Valente, H. Failache, A. Lezama, Polarization squeezing of light by single passage through an atomic vapor. *Phys. Rev. A* **84**, 033851 (2011)
151. J. Ries, B. Brezger, A.I. Lvovsky, Experimental vacuum squeezing in rubidium vapor via self-rotation. *Phys. Rev. A* **68**, 025801 (2003)
152. K.M. Birnbaum, A. Boca, R. Miller, A.D. Boozer, T.E. Northup, H.J. Kimble, Photon blockade in an optical cavity with one trapped atom. *Nature* **436**, 87–90 (2005)
153. B. Dayan, A.S. Parkins, T. Aoki, E.P. Ostby, K.J. Vahala, H.J. Kimble, A photon turnstile dynamically regulated by one atom. *Science* **319**, 1062–1065 (2008)
154. C.S. Muñoz, E. del Valle, A.G. Tudela, K. Müller, S. Lichtmannecker, M. Kaniber, C. Tejedor, J.J. Finley, F.P. Laussy, Emitters of n-photon bundles. *Nat. Phot.* **8**, 550–555 (2014)
155. D.V. Strekalov, A bundle of photons, please. *Nat. Phot.* **8**, 500–501 (2014)
156. T. Basche, W.E. Moerner, M. Orrit, H. Talon, Photon antibunching in the fluorescence of a single dye molecule trapped in a solid. *Phys. Rev. Lett.* **69**, 1516–1519 (1992)
157. C. Brunel, B. Lounis, P. Tamarat, M. Orrit, Triggered source of single photons based on controlled single molecule fluorescence. *Phys. Rev. Lett.* **83**, 2722–2725 (1999)

158. B. Lounis, W.E. Moerner, Single photons on demand from a singlemolecule at room temperature. *Nature* **407**, 491–493 (2000)
159. B. Lounis, M. Orrit, Single-photon sources. *Rep. Prog. Phys.* **68**, 1129–1179 (2005)
160. S. Buckley, K. Rivoire, J. Vučković, Engineered quantum dot single-photon sources. *Rep. Prog. Phys.* **75**, 126503 (2012)
161. P. Michler, A. Imamoglu, M.D. Maso, P.J. Carson, G.F. Strouse, S.K. Buratto, Quantum correlation among photons from a single quantum dot at room temperature. *Nature* **406**, 968–970 (2000)
162. S. Bounouar, M. Elouneig-Jamroz, M. d. Hertog, C. Morschut, E. Bellet-Amalric, R. André, C. Bougerol, Y. Genuist, J.-P. Poizat, S. Tatarenko, K. Kheng, Ultrafast room temperature single-photon source from nanowire-quantum dots. *Nano Lett.* **12**, 2977–2981 (2012)
163. M.J. Holmes, K. Choi, S. Kako, M. Arita, Y. Arakawa, Room-temperature triggered single photon emission from a iii-nitride site-controlled nanowire quantum dot. *Nano Lett.* **14**, 982–986 (2014)
164. A. Högele, C. Galland, M. Winger, A. Imamolu, Photon antibunching in the photoluminescence spectra of a single carbon nanotube. *Phys. Rev. Lett.* **100**, 217401 (2008)
165. S. Schietinger, T. Schröder, O. Benson, One-by-one coupling of single defect centers in nanodiamonds to high-q modes of an optical microresonator. *Nano Lett.* **8**(11), 3911–3915 (2008)
166. T.M. Babinec, B.J.M. Hausmann, M. Khan, Y. Zhang, J.R. Maze, P.R. Hemmer, M. Loncar, A diamond nanowire single-photon source. *Nat. Nanotech.* **5**, 195–199 (2010)
167. C.H.H. Schulte, J. Hansom, A.E. Jones, C. Matthiesen, C. Le Gall, M. Atatüre, Quadrature squeezed photons from a two-level system. *Nature* **525**, 222–225 (2015)
168. D. Press, S. Götzinger, S. Reitzenstein, C. Hofmann, A. Löffler, M. Kamp, A. Forchel, Y. Yamamoto, Photon antibunching from a single quantum-dot-microcavity system in the strong coupling regime. *Phys. Rev. Lett.* **98**, 117402 (2007)
169. S. Strauf, N.G. Stoltz, M.T. Rakher, L.A. Coldren, P.M. Petroff, D. Bouwmeester, High-frequency single-photon source with polarization control. *Nat. Phot.* **1**, 704–708 (2007)
170. E. Peter, P. Senellart, D. Martrou, A. Lemaître, J. Hours, J.M. Gérard, J. Bloch, Exciton-photon strong-coupling regime for a single quantum dot embedded in a microcavity. *Phys. Rev. Lett.* **95**, 067401 (2005)
171. K. Srinivasan, O. Painter, Linear and nonlinear optical spectroscopy of a strongly coupled microdisk-quantum dot system. *Nature* **450**, 862–865 (2007)
172. M.N. Makhonin, J.E. Dixon, R.J. Coles, B. Royall, I.J. Luxmoore, E. Clarke, M. Hugues, M.S. Skolnick, A.M. Fox, Waveguide coupled resonance fluorescence from on-chip quantum emitter. *Nano Lett.* **14**, 6997–7002 (2014)
173. N. Akopian, N.H. Lindner, E. Poem, Y. Berlatzky, J. Avron, D. Gershoni, B.D. Gerardot, P.M. Petroff, Entangled photon pairs from semiconductor quantum dots. *Phys. Rev. Lett.* **96**, 130501 (2006)
174. R.M. Stevenson, R.J. Young, P. Atkinson, K. Cooper, D.A. Ritchie, A.J. Shields, A semiconductor source of triggered entangled photon pairs. *Nature* **439**, 179–182 (2006)
175. T. Kuroda, T. Mano, N. Ha, H. Nakajima, H. Kumano, B. Urbaszek, M. Jo, M. Abbarchi, Y. Sakuma, K. Sakoda, I. Suemune, X. Marie, T. Amand, Symmetric quantum dots as efficient sources of highly entangled photons: Violation of bell's inequality without spectral and temporal filtering. *Phys. Rev. B* **88**, 041306 (2013)
176. H. Jayakumar, A. Predojević, T. Huber, T. Kauten, G.S. Solomon, G. Weihs, Deterministic photon pairs and coherent optical control of a single quantum dot. *Phys. Rev. Lett.* **110**, 135505 (2013)
177. N. Dotti, F. Sarti, S. Bietti, A. Azarov, A. Kuznetsov, F. Biccari, A. Vinattieri, S. Sanguinetti, M. Abbarchi, M. Gurioli, Germanium-based quantum emitters towards a time-reordering entanglement scheme with degenerate exciton and biexciton states. *Phys. Rev. B* **91**, 205316 (2015)
178. R. Trotta, J.S. Wildmann, E. Zallo, O.G. Schmidt, A. Rastelli, Highly entangled photons from hybrid piezoelectric-semiconductor quantum dot devices. *Nano Lett.* **14**, 3439–3444 (2014)

179. R.J. Young, R.M. Stevenson, P. Atkinson, K. Cooper, D.A. Ritchie, A.J. Shields, Improved fidelity of triggered entangled photons from single quantum dots. *New J. Phys.* **8**, 29 (2006)
180. M. Müller, S. Bounouar, K.D. Jöns, M. Glässl, P. Michler, On-demand generation of indistinguishable polarization-entangled photon pairs. *Nat. Phot.* **8**, 224–228 (2014)
181. J.A. Giordmaine, R.C. Miller, Tunable coherent parametric oscillation in LiNbO₃ at optical frequencies. *Phys. Rev. Lett.* **14**, 973–976 (1965)
182. D.N. Klyshko, Scattering of light in a medium with nonlinear polarizability. *JETP Lett.* **28**, 522–526 (1969)
183. D. Strekalov, A.B. Matsko, A.A. Savchenkov, L. Maleki, Relationship between quantum two-photon correlation and classical spectrum of light. *Phys. Rev. A* **71**, 041803 (2005)
184. M.H. Rubin, D.N. Klyshko, Y.H. Shih, A.V. Sergienko, Theory of two-photon entanglement in type-II optical parametric down-conversion. *PRA* **50**, 5122–5133 (1994)
185. E. Dauler, G. Jaeger, A. Muller, A. Migdall, A. Sergienko, Tests of a two-photon technique for measuring polarization mode dispersion with subfemtosecond precision. *J. Res. Natl. Inst. Stand. Technol.* **104**, 1–10 (1999)
186. A. Valencia, M.V. Chekhova, A. Trifonov, Y. Shih, Entangled two-photon wave packet in a dispersive medium. *Phys. Rev. Lett.* **88**, 183601 (2002)
187. D. Strekalov, A.B. Matsko, A. Savchenkov, L. Maleki, Quantum-correlation metrology with biphotons: where is the limit? *J. Mod. Opt.* **52**, 2233–2243 (2005)
188. M. Scholz, L. Koch, O. Benson, Statistics of narrow-band single photons for quantum memories generated by ultrabright cavity-enhanced parametric down-conversion. *Phys. Rev. Lett.* **102**, 63603 (2009)
189. C.-S. Chuu, G.Y. Yin, S.E. Harris, A miniature ultrabright source of temporally long, narrowband biphotons. *Appl. Phys. Lett.* **101**, 051108 (2012)
190. M. Förtsch, J.U. Fürst, C. Wittmann, D. Strekalov, A. Aiello, M.V. Chekhova, C. Silberhorn, G. Leuchs, C. Marquardt, A versatile source of single photons for quantum information processing. *Nat. Commun.* **4**, 1818 (2013)
191. A.V. Burlakov, M.V. Chekhova, D.N. Klyshko, S.P. Kulik, A.N. Penin, Y.H. Shih, D.V. Strekalov, Interference effects in spontaneous two-photon parametric scattering from two macroscopic regions. *Phys. Rev. A* **56**, 3214–3225 (1997)
192. T.S. Iskhakov, S. Lemieux, A. Perez, R.W. Boyd, G. Leuchs, M.V. Chekhova, Nonlinear interferometer for tailoring the frequency spectrum of bright squeezed vacuum. *J. Mod. Opt.* **63**, 64–70 (2016)
193. T. Setälä, T. Shirai, A.T. Friberg, Fractional fourier transform in temporal ghost imaging with classical light. *Phys. Rev. A* **82**, 043813 (2010)
194. D. Sych, V. Averchenko, G. Leuchs, Shaping a single photon without interacting with it. *Phys. Rev. A* **96**, 053847 (2017)
195. V. Averchenko, D. Sych, G. Leuchs, Heralded temporal shaping of single photons enabled by entanglement. *Phys. Rev. A* **96**, 043822 (2017)
196. P.R. Tapster, J.G. Rarity, Photon statistics of pulsed parametric light. *J. Mod. Opt.* **45**, 595–604 (1998)
197. P.G. Kwiat, E. Waks, A.G. White, I. Appelbaum, P.H. Eberhard, Ultrabright source of polarization-entangled photons. *Phys. Rev. A* **60**, R773–R776 (1999)
198. J.T. Barreiro, N.K. Langford, N.A. Peters, P.G. Kwiat, Generation of hyperentangled photon pairs. *Phys. Rev. Lett.* **95**, 260501 (2005)
199. T.S. Iskhakov, A.M. Pérez, K.Y. Spasibko, M.V. Chekhova, G. Leuchs, Superbunched bright squeezed vacuum state. *Opt. Lett.* **37**, 1919–1921 (2012)
200. K. Sanaka, K. Kawahara, T. Kuga, New high-efficiency source of photon pairs for engineering quantum entanglement. *Phys. Rev. Lett.* **86**, 5620–5623 (2001)
201. G. Harder, V. Ansari, B. Brecht, T. Dürmeier, C. Marquardt, C. Silberhorn, An optimized photon pair source for quantum circuits. *Opt Express* **21**, 13975–13985 (2013)
202. A.M. Pérez, K.Y. Spasibko, P.R. Sharapova, O.V. Tikhonova, G. Leuchs, M.V. Chekhova, Giant narrowband twin-beam generation along the pump-energy propagation direction. *Nat. Commun.* **6**, 7707 (2015)

203. M. Zukowski, A. Zeilinger, M.A. Horne, A.K. Ekert, Event-ready-detectors bell experiment via entanglement swapping. *Phys. Rev. Lett.* **71**, 4287–4290 (1993)
204. M. Zukowski, A. Zeilinger, H. Weinfurter, Entangling photons radiated by independent pulsed sources. *Ann. NY Acad. Sci.* **755**, 91 (1995)
205. M. R  dmark, M. Zukowski, M. Bourennane, Experimental test of fidelity limits in six-photon interferometry and of rotational invariance properties of the photonic six-qubit entanglement singlet state. *Phys. Rev. Lett.* **103**, 150501 (2009)
206. O. Aytur, P. Kumar, Squeezed-light generation with a mode-locked q-switched laser and detection by using a matched local oscillator. *Opt. Lett.* **17**, 529–531 (1992)
207. C. Kim, P. Kumar, Quadrature-squeezed light detection using a self-generated matched local oscillator. *Phys. Rev. Lett.* **73**, 1605–1608 (1994)
208. K. Hirosawa, Y. Ito, H. Ushio, H. Nakagome, F. Kannari, Generation of squeezed vacuum pulses using cascaded second-order optical nonlinearity of periodically poled lithium niobate in a sagnac interferometer. *Phys. Rev. A* **80**, 043832 (2009)
209. M. Pysher, R. Bloomer, C.M. Kaleva, T.D. Roberts, B. Philip, O. Pfister, Broadband amplitude squeezing in a periodically poled KTiOPO₄ waveguide. *Opt. Lett.* **34**, 256–258 (2009)
210. G. Breitenbach, S. Schiller, J. Mlynek, Measurement of the quantum states of squeezed light. *Nature* **387**, 471–475 (1997)
211. M. Lassen, M. Sabuncu, P. Buchhave, U.L. Andersen, Generation of polarization squeezing with periodically poled KTP at 1064 nm. *Opt. Expr.* **15**, 5077–5082 (2007)
212. A.M. P  rez, T.S. Iskhakov, P. Sharapova, S. Lemieux, O.V. Tikhonova, M.V. Chekhova, G. Leuchs, Bright squeezed-vacuum source with 1.1 spatial mode. *Opt. Lett.* **39**, 2403–2406 (2014)
213. Z. Yan, X. Jia, X. Su, Z. Duan, C. Xie, K. Peng, Cascaded entanglement enhancement. *Phys. Rev. A* **85**, 040305 (2012)
214. L.-A. Wu, H.J. Kimble, J.L. Hall, H. Wu, Generation of squeezed states by parametric down conversion. *Phys. Rev. Lett.* **57**, 2520–2523 (1986)
215. S. Suzuki, H. Yonezawa, F. Kannari, M. Sasaki, A. Furusawa, 7 dB quadrature squeezing at 860 nm with periodically poled KTiOPO₄. *Appl. Phys. Lett.* **89**, 061116 (2006)
216. Y. Takeno, M. Yukawa, H. Yonezawa, A. Furusawa, Observation of –9 dB quadrature squeezing with improvement of phase stability in homodyne measurement. *Opt. Express* **15**, 4321–4327 (2007)
217. G. H  tet, O. Gl  ckl, K.A. Pilypas, C.C. Harb, B.C. Buchler, H.-A. Bachor, P.K. Lam, Squeezed light for bandwidth-limited atom optics experiments at the rubidium d1 line. *J. Phys. B: At. Mol. Opt. Phys.* **40**, 221–226 (2007)
218. H. Vahlbruch, M. Mehmet, S. Chelkowski, B. Hage, A. Franzen, N. Lastzka, S. Gossler, K. Danzmann, R. Schnabel, Observation of squeezed light with 10-dB quantum-noise reduction. *Phys. Rev. Lett.* **100**, 033602 (2008)
219. T. Eberle, S. Steinlechner, J. Bauchrowitz, V. H  ndchen, H. Vahlbruch, M. Mehmet, H. M  ller-Ebhardt, R. Schnabel, Quantum enhancement of the zero-area sagnac interferometer topology for gravitational wave detection. *Phys. Rev. Lett.* **104**, 251102 (2010)
220. A. Heidmann, R.J. Horowicz, S. Reynaud, E. Giacobino, C. Fabre, G. Camy, Observation of quantum noise reduction on twin laser beams. *Phys. Rev. Lett.* **59**, 2555–2557 (1987)
221. J. Mertz, T. Debuisschert, A. Heidmann, C. Fabre, E. Giacobino, Improvements in the observed intensity correlation of optical parametric oscillator twin beams. *Opt. Lett.* **16**, 1234–1236 (1991)
222. P.R. Tapster, J.G. Rarity, J.S. Satchell, Use of parametric down-conversion to generate sub-poissonian light. *Phys. Rev. A* **37**, 2963–2967 (1988)
223. J. Mertz, A. Heidmann, C. Fabre, Generation of sub-poissonian light using active control with twin beams. *Phys. Rev. A* **44**, 3229–3238 (1991)
224. J. Laurat, T. Coudreau, N. Treps, A. Ma  tre, C. Fabre, Conditional preparation of a quantum state in the continuous variable regime: generation of a sub-poissonian state from twin beams. *Phys. Rev. Lett.* **91**, 213601 (2003)

225. J. Hald, J.L. Sørensen, C. Schori, E.S. Polzik, Spin squeezed atoms: a macroscopic entangled ensemble created by light. *Phys. Rev. Lett.* **83**, 1319–1322 (1999)
226. K. Honda, D. Akamatsu, M. Arikawa, Y. Yokoi, K. Akiba, S. Nagatsuka, T. Tanimura, A. Furusawa, M. Kozuma, Storage and retrieval of a squeezed vacuum. *Phys. Rev. Lett.* **100**, 093601 (2008)
227. M. Scholz, L. Koch, R. Ullmann, O. Benson, Single-mode operation of a high-brightness narrow-band single-photon source. *Appl. Phys. Lett.* **94**, 201105 (2009)
228. F. Wolfgramm, Y.A. de Icaza Astiz, F.A. Beduini, A. Cerè, M.W. Mitchell, Atom-resonant heralded single photons by interaction-free measurement. *Phys. Rev. Lett.* **106**, 053602 (2011)
229. A.B. Matsko, V.S. Ilchenko, Optical resonators with whispering-gallery modes-part I: basics. *J. Sel. Top. Quantum Electron.* **12**, 3 (2006)
230. V.S. Ilchenko, A.B. Matsko, Optical resonators with whispering-gallery modes-part II: applications. *J. Sel. Top. Quantum Electron.* **12**, 15–32 (2006)
231. A. Chiasera, Y. Dumeige, P. Féron, M. Ferrari, Y. Jestin, G. Nunzi Conti, S. Pelli, S. Soria, G.C. Righini, Spherical whispering-gallery-mode microresonators. *Las. Phot. Rev.* **4**, 457–482 (2010)
232. D.V. Strekalov, C. Marquardt, A.B. Matsko, H.G.L. Schwefel, G. Leuchs, Nonlinear and quantum optics with whispering gallery resonators. *J. Opt.* **18**, 123002 (2016)
233. D.V. Strekalov, A.S. Kowlgy, Y.-P. Huang, P. Kumar, Optical sum-frequency generation in a whispering-gallery-mode resonator. *New J. Phys.* **16**, 053025 (2014)
234. A.A. Savchenkov, A.B. Matsko, M. Mohageg, D.V. Strekalov, L. Maleki, Parametric oscillations in a whispering gallery resonator. *Opt. Lett.* **32**, 157–159 (2007)
235. J.U. Fürst, D.V. Strekalov, D. Elser, A. Aiello, U.L. Andersen, C. Marquardt, G. Leuchs, Low-threshold optical parametric oscillations in a whispering gallery mode resonator. *Phys. Rev. Lett.* **105**, 263904 (2010)
236. T. Beckmann, H. Linnenbank, H. Steigerwald, B. Sturman, D. Haertle, K. Buse, I. Breunig, Highly tunable low-threshold optical parametric oscillation in radially poled whispering gallery resonators. *Phys. Rev. Lett.* **106**, 143903 (2011)
237. T. Beckmann, K. Buse, I. Breunig, Optimizing pump threshold and conversion efficiency of whispering gallery optical parametric oscillators by controlled coupling. *Opt. Lett.* **37**, 5250–5252 (2012)
238. C.S. Werner, T. Beckmann, K. Buse, I. Breunig, Blue-pumped whispering gallery optical parametric oscillator. *Opt. Lett.* **37**, 4224–4226 (2012)
239. C.S. Werner, K. Buse, I. Breunig, Continuous-wave whispering-gallery optical parametric oscillator for high-resolution spectroscopy. *Opt. Lett.* **40**, 772–775 (2015)
240. M. Förtsch, T. Gerrits, M.J. Stevens, D. Strekalov, G. Schunk, J.U. Fürst, U. Vogl, F. Sedlmeir, H.G.L. Schwefel, G. Leuchs, S.W. Nam, C. Marquardt, Near-infrared single-photon spectroscopy of a whispering gallery mode resonator using energy-resolving transition edge sensors. *J. Opt.* **17**, 065501 (2015)
241. A. Sizmann, R.J. Horowitz, G. Wagner, G. Leuchs, Observation of amplitude squeezing of the up-converted mode in second harmonic generation. *Opt. Commun.* **80**, 138–142 (1990)
242. P. Kurz, R. Paschotta, K. Fiedler, A. Sizmann, G. Leuchs, J. Mlynek, Squeezing by second-harmonic generation in a monolithic resonator. *Appl. Phys. B* **55**, 216–225 (1992)
243. P.D. Drummond, K.J. McNeil, D.F. Walls, Non-equilibrium transitions in sub/second harmonic generation II: quantum theory. *Opt. Acta* **28**, 211–225 (1981)
244. S.F. Pereira, M. Xiao, H.J. Kimble, J.L. Hall, Generation of squeezed light by intracavity frequency doubling. *Phys. Rev. A* **38**, 4931 (1988)
245. B. Hage, A. Samblowski, R. Schnabel, Towards einstein-podolsky-rosen quantum channel multiplexing. *Phys. Rev. A* **81**, 062301 (2010)
246. M. Pysher, Y. Miwa, R. Shahrokhshahi, R. Bloomer, O. Pfister, Parallel generation of quadruparticle cluster entanglement in the optical frequency comb. *Phys. Rev. Lett.* **107**, 030505 (2011)
247. A. Brieussel, Y. Shen, G. Campbell, G. Guccione, J. Janousek, B. Hage, B.C. Buchler, N. Treps, C. Fabre, F.Z. Fang, X.Y. Li, T. Symul, P.K. Lam, Squeezed light from a diamond-turned monolithic cavity. *Opt. Express* **24**, 4042 (2016)

248. C.C. Gerry, P.L. Knight, *Introductory Quantum Optics* (Cambridge University Press, 2005)
249. R.Y. Chiao, E. Garmire, C.H. Townes, Self-trapping of optical beams. *Phys. Rev. Lett.* **13**, 479–482 (1964)
250. M. Kitagawa, Y. Yamamoto, Number-phase minimum-uncertainty state with reduced number uncertainty in a kerr nonlinear interferometer. *Phys. Rev. A* **34**, 3974–3988 (1986)
251. R.M. Shelby, M.D. Levenson, S.H. Perlmutter, R.G. DeVoe, D.F. Walls, Broad-band parametric deamplification of quantum noise in an optical fiber. *Phys. Rev. Lett.* **57**, 691–694 (1986)
252. K. Bergman, H.A. Haus, Squeezing in fibers with optical pulses. *Opt. Lett.* **16**, 663–665 (1991)
253. M. Rosenbluh, R.M. Shelby, Squeezed optical solitons. *Phys. Rev. Lett.* **66**, 153–156 (1991)
254. S. Schmitt, J. Ficker, M. Wolff, F. König, A. Sizmann, G. Leuchs, Photon-number squeezed solitons from an asymmetric fiber-optic sagnac interferometer. *Phys. Rev. Lett.* **81**, 2446–2449 (1998)
255. D. Krylov, K. Bergman, Amplitude-squeezed solitons from an asymmetric fiber interferometer. *Opt. Lett.* **23**, 1390–1392 (1998)
256. S.R. Friberg, S. Machida, M.J. Werner, A. Levanon, T. Mukai, Observation of optical soliton photon-number squeezing. *Phys. Rev. Lett.* **77**, 3775–3778 (1996)
257. S. Spälter, M. Burk, U. Strößner, A. Sizmann, G. Leuchs, Propagation of quantum properties of subpicosecond solitons in a fiber. *Opt. Express* **2**, 77–83 (1998)
258. C. Riek, P. Sulzer, M. Seeger, A.S. Moskalenko, G. Burkard, D.V. Seletskiy, A. Leitenstorfer, Subcycle quantum electrodynamics. *Nature* **541**, 376–379 (2017)
259. A.S. Moskalenko, C. Riek, D.V. Seletskiy, G. Burkard, A. Leitenstorfer, Paraxial theory of direct electro-optic sampling of the quantum vacuum. *Phys. Rev. Lett.* **115**, 263601 (2015)
260. C. Riek, D.V. Seletskiy, A.S. Moskalenko, J.F. Schmidt, P. Krauspe, S. Eckart, S. Eggert, G. Burkard, A. Leitenstorfer, Direct sampling of electric-field vacuum fluctuations. *Science* **350**, 420–423 (2015)
261. D. Levandovsky, M. Vasilyev, P. Kumar, Amplitude squeezing of light by means of a phase-sensitive fiber parametric amplifier. *Opt. Lett.* **24**, 984–986 (1999)
262. J. Heersink, V. Josse, G. Leuchs, U.L. Andersen, Efficient polarization squeezing in optical fibers. *Opt. Lett.* **30**, 1192–1194 (2005)
263. R. Dong, J. Heersink, J.F. Corney, P.D. Drummond, U.L. Andersen, G. Leuchs, Experimental evidence for raman-induced limits to efficient squeezing in optical fibers. *Opt. Lett.* **33**, 116–118 (2008)
264. M. Margalit, C.X. Xu, E.P. Ippen, H.A. Haus, Cross phase modulation squeezing in optical fibers. *Opt. Express* **2**, 72–76 (1998)
265. K. Hirosawa, H. Furumochi, A. Tada, F. Kannari, M. Takeoka, M. Sasaki, Photon number squeezing of ultrabroadband laser pulses generated by microstructure fibers. *Phys. Rev. Lett.* **94**, 203601 (2005)
266. J. Milanovic, M. Lassen, U.L. Andersen, G. Leuchs, A novel method for polarization squeezing with photonic crystal fibers. *Opt. Express* **18**, 1521–1527 (2010)
267. J.G. Rarity, J. Fulconis, J. Dulingall, W.J. Wadsworth, P.S.J. Russell, Photonic crystal fiber source of correlated photon pairs. *Opt. Express* **13**, 534–544 (2005)
268. J. Fan, A. Migdall, A broadband high spectral brightness fiberbased two-photon source. *Opt. Express* **15**, 2915–2920 (2007)
269. J. Nold, P. Hölzer, N.Y. Joly, G.K.L. Wong, A. Nazarkin, A. Podlipensky, M. Scharrer, P.S.J. Russell, Pressure-controlled phase matching to third harmonic in ar-filled hollow-core photonic crystal fiber. *Opt. Lett.* **35**, 2922–2924 (2010)
270. M.A. Finger, T.S. Iskhakov, N.Y. Joly, M.V. Chekhova, P.S.J. Russell, Raman-free, noble-gas-filled photonic-crystal fiber source for ultrafast, very bright twin-beam squeezed vacuum. *Phys. Rev. Lett.* **115**, 143602 (2015)
271. U. Vogl, N.Y. Joly, P.S.J. Russell, C. Marquardt, G. Leuchs, Squeezed light and self-induced transparency in mercury-filled hollow core photonic crystal fibers (2015)

272. T.D. Bradley, Y. Wang, M. Alharbi, B. Debord, C. Fourcade-Dutin, B. Beaudou, F. Gerome, F. Benabid, Optical properties of low loss (70 dB/km) hypocycloid-core Kagome hollow core photonic crystal fiber for Rb and Cs based optical applications. *J. Lightwave Tech.* **31**, 2752–2755 (2013)
273. Y.K. Chembo, D.V. Strekalov, N. Yu, Spectrum and dynamics of optical frequency combs generated with monolithic whispering gallery mode resonators. *Phys. Rev. Lett.* **104**, 103902 (2010)
274. P. Del'Haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, T.J. Kippenberg, Optical frequency comb generation from a monolithic microresonator. *Nature* **450**, 1214–1217 (2007)
275. W. Liang, A.A. Savchenkov, Z. Xie, J.F. McMillan, J. Burkhart, V.S. Ilchenko, C.W. Wong, A.B. Matsko, L. Maleki, Miniature multi octave light source based on a monolithic microcavity. *Optica* **2**, 40 (2015)
276. S. Clemmen, K.P. Huy, W. Bogaerts, R.G. Baets, P. Emplit, S. Massar, Continuous wave photon pair generation in silicon-on-insulator waveguides and ring resonators. *Opt. Express* **17**(19), 16558–16570 (2009)
277. S. Azzini, D. Grassani, M.J. Strain, M. Sorel, L.G. Helt, J.E. Sipe, M. Liscidini, M. Galli, D. Bajoni, Ultra-low power generation of twin photons in a compact silicon ring resonator. *Opt. Express* **20**(21), 23100–23107 (2012)
278. E. Engin, D. Bonneau, C.M. Natarajan, A.S. Clark, M.G. Tanner, R.H. Hadfield, S.N. Dorenbos, V. Zwiller, K. Ohira, N. Suzuki, H. Yoshida, N. Iizuka, M. Ezaki, J.L. O'Brien, M.G. Thompson, Photon pair generation in a silicon micro-ring resonator with reverse bias enhancement. *Opt. Express* **21**(23), 27826 (2013)
279. Y. Guo, W. Zhang, S. Dong, Y. Huang, J. Peng, Telecom-band degenerate-frequency photon pair generation in silicon microring cavities. *Opt. Lett.* **39**(8), 2526–2529 (2014)
280. D. Grassani, S. Azzini, M. Liscidini, M. Galli, M.J. Strain, M. Sorel, J.E. Sipe, D. Bajoni, Micrometer-scale integrated silicon source of time-energy entangled photons. *Optica* **2**(2), 88–94 (2015)
281. R. Wakabayashi, M. Fujiwara, K.-I. Yoshino, Y. Nambu, M. Sasaki, T. Aoki, Time-bin entangled photon pair generation from Si micro-ring resonator. *Opt. Express* **23**(2), 1103 (2015)
282. J. Suo, S. Dong, W. Zhang, Y. Huang, J. Peng, Generation of hyper-entanglement on polarization and energy-time based on a silicon micro-ring cavity. *Opt. Express* **23**(4), 3985–3995 (2015)
283. A. Dutt, K. Luke, S. Manipatruni, A.L. Gaeta, P. Nussenzveig, M. Lipson, On-chip optical squeezing. *Phys. Rev. Appl.* **3**, 044005 (2015)
284. U.B. Hoff, B.M. Nielsen, U.L. Andersen, Integrated source of broadband quadrature squeezed light. *Opt. Express* **23**, 12013–12036 (2015)
285. T.P. Purdy, P.-L. Yu, R.W. Peterson, N.S. Kampel, C.A. Regal, Strong optomechanical squeezing of light. *Phys. Rev. X* **3**, 031012 (2013)
286. A.H. Safavi-Naeini, S. Gröblacher, J.T. Hill, J. Chan, M. Aspelmeyer, O. Painter, Squeezed light from a silicon micromechanical resonator. *Nature* **500**, 185–189 (2013)
287. M.C. Teich, B.E.A. Saleh, Observation of sub-poisson Franck-hertz light at 253.7 nm. *JOSA B* **2**, 275–282 (1985)
288. W. Schottky, E. Spehnke, Raumladungsschwächung des schroeffekts. *Wiss. Veröff. Siemens-Werke* **16**, 1–18 (1937)
289. Y. Yamamoto, S. Machida, High-impedance suppression of pump fluctuation and amplitude squeezing. *Phys. Rev. A* **35**, 5114–5130 (1987)
290. S. Machida, Y. Yamamoto, Y. Itaya, Observation of amplitude squeezing in a constant-current- driven semiconductor laser. *Phys. Rev. Lett.* **58**, 1000–1003 (1987)
291. S. Machida, Y. Yamamoto, Ultrabroadband amplitude squeezing in a semiconductor laser. *Phys. Rev. Lett.* **60**, 792–794 (1988)
292. W.H. Richardson, S. Machida, Y. Yamamoto, Squeezed photon-number noise and sub-poissonian electrical partition noise in a semiconductor laser. *Phys. Rev. Lett.* **66**, 2867–2870 (1991)

293. F. Marin, A. Bramati, E. Giacobino, T.-C. Zhang, J.P. Poizat, J.-F. Roch, P. Grangier, Squeezing and intermode correlations in laser diodes. *Phys. Rev. Lett.* **75**, 4606–4609 (1995)
294. I. Maurin, I. Protsenko, J.-P. Hermier, A. Bramati, P. Grangier, E. Giacobino, Light intensity-voltage correlations and leakage-current excess noise in a single-mode semiconductor laser. *Phys. Rev. A* **72**, 033823 (2005)
295. H. Wang, M.J. Freeman, D.G. Steel, Squeezed light from injection-locked quantum well lasers. *Phys. Rev. Lett.* **71**, 3951–3954 (1993)
296. M.J. Freeman, H. Wang, D.G. Steel, R. Craig, D.R. Scifres, Wavelength-tunable amplitude-squeezed light from a room-temperature quantum-well laser. *Opt. Lett.* **18**, 2141–2143 (1993)
297. F. Wolf, R.G. Ispasoiu, J.F. Ryan, A.M. Fox, Photon-number squeezing in a free-running quantum-well laser operating at 980 nm. *J. Opt. B: Quantum Semiclass. Opt.* **4**, 129–133 (2002)
298. M. Uemukai, S. Nozu, T. Suhara, High-efficiency InGaAs QW distributed bragg reflector laser with curved grating for squeezed light generation. *J. Sel. Top. Quantum Electron.* **11**, 1143–1147 (2005)
299. Y. Yamamoto, N. Imoto, S. Machida, Amplitude squeezing in a semiconductor laser using quantum nondemolition measurement and negative feedback. *Phys. Rev. A* **33**, 3243–3261 (1986)
300. B.C. Buchler, M.B. Gray, D.A. Shaddock, T.C. Ralph, D.E. McClelland, Suppression of classic and quantum radiation pressure noise by electro-optic feedback. *Opt. Lett.* **24**, 259–261 (1999)
301. J.H. Shapiro, G. Saplakoglu, S.-T. Ho, P. Kumar, B.E.A. Saleh, M.C. Teich, Theory of light detection in the presence of feedback. *JOSA B* **4**, 1604–1620 (1987)
302. S. Mancini, D. Vitali, P. Tombesi, Motional squashed states. *J. Opt. B: Quantum Semiclass. Opt.* **2**, 190–195 (2000)
303. A.O. Caldeira, A.J. Leggett, Influence of damping on quantum interference: an exactly soluble model. *Phys. Rev. A* **31**, 1059–1066 (1985)
304. G. Leuchs, U. Andersen, The effect of dissipation on non-classical states of the radiation field. *Las. Phys.* **15**, 129–134 (2005)
305. J.H. Eberly, N.B. Narozhny, J.J. Sanchez-Mondragon, Periodic spontaneous collapse and revival in a simple quantum model. *Phys. Rev. Lett.* **44**, 1323–1326 (1980)
306. G. Rempe, H. Walther, N. Klein, Observation of quantum collapse and revival in a one-atom maser. *Phys. Rev. Lett.* **58**, 353–356 (1987)

Chapter 4

Nanophotonic Advances for Room-Temperature Single-Photon Sources



Svetlana G. Lukishova and Luke J. Bissell

Abstract This review reports on recent advances in room-temperature single-photon sources (SPSs) with photons exhibiting antibunching (separation of all photons in time in contrast to faint laser sources), including the authors' results on SPSs with definite circular and linear polarizations. SPSs are important devices in secure quantum communication. Some quantum computing schemes are also based on such sources. Quantum metrology, quantum memory and fundamental physics experiments are other applications of SPSs. The critical issue in producing “antibunched” photons is the very low concentration of photon emitters, such that, within an excitation-laser focal spot, only one emitter becomes excited and which will emit only one photon at a time. Single “giant” colloidal semiconductor nanocrystal quantum dots and dot-in-rods, diamond color centers (both bulk and nanodiamonds), and trivalent rare-earth ions (TR^{3+}) have the best photostability (longest operating time) in room-temperature excitation. This review is focused on nanophotonic aspects of the problem, describing room-temperature SPSs based on these emitters and some new stable single-emitters. We also describe methods for emitter fluorescence enhancement: microcavities (including photonic bandgap, Bragg reflector and chiral liquid crystal microcavities), plasmonic nanoantennas, and metamaterials. Finally, we describe the alignment of anisotropic single emitters with liquid crystals.

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road, Wilmot Building, Rochester, NY 14627, USA

e-mail: sruk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

L. J. Bissell (✉)

Air Force Research Laboratory, Materials and Manufacturing Directorate, 2179 12th St., WPAFB, OH 45433, USA

e-mail: luke.bissell@us.af.mil

4.1 Introduction

Single-photon sources (SPSs), producing nonclassical light, are pivotal components in quantum communication technology [1, 2]. Ideal SPSs efficiently produce individual photons separated in time with no chance for overlap of photons (anti-bunching) [3–5]. For common light sources (incandescent bulbs, sun light) and in lasers, multiple (nonseparated) photons cannot be avoided even in very faint sources. Using photons that are all temporally separated, secure quantum communication will prevent any potential eavesdropper (Eve) from intercepting a message without the receiver (Bob) noticing it. In the popular BB84 quantum key distribution (QKD) protocol [6] bits of transferred information between a sender (Alice) and the receiver are coded in the photon polarization state. If the source emits more than one photon at a time, it would be possible for Eve to extract one of these photons without disturbing the other photons. This is the so-called beam-splitter attack which does not reveal Eve's presence. Although it was proved that for very faint classical light sources, e.g., with mean photon number less than 0.1, the security of the BB84 protocol is just as good with faint laser pulses as with SPSs [7], the price to pay for using such faint classical sources is a significant reduction in bit rate [1, 2]. In the decoy state approach [8, 9] Alice modulates a weak classical signal and can achieve the same security, as ideal SPSs, albeit with a lower bit rate. However due to losses for long distance QKD quantum repeaters based on single-photon sources become necessary ([10–14], see also [15–18] on quantum memory). In another implementation, a SPS becomes the critical hardware element for quantum computers with linear optical elements and photodetectors [19–22]. In addition, single photons are promising in quantum metrology [23, 24] and fundamental physics experiments, such as in interference from uncoupled light sources [25] and entanglement-state generation [26], in shedding light on wave-particle duality as well as in fundamental issues of quantum measurements and uncertainty.

4.1.1 Photon Statistics of Different Light Sources

In this section, we provide only a brief overview of photon statistics. For more details see [27]. Figure 4.1 illustrates the difference in separation in time between

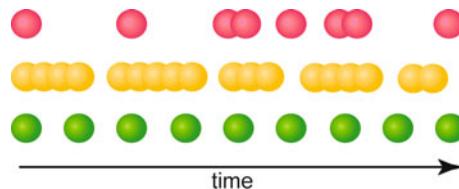


Fig. 4.1 Photon separation in time for sources with different photon statistics: top (red color)—coherent (laser) light, center (yellow color)—thermal light (bunching), bottom (green color)—single (antibunched) photons

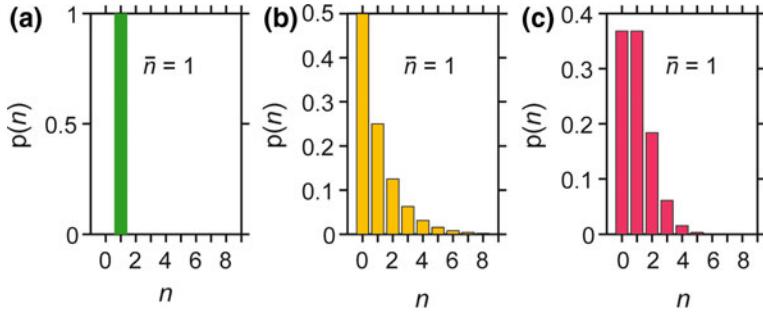


Fig. 4.2 Probability distribution [$P(n)$] versus photon number (n) for photon sources with different statistics, but with the same mean photon number $\bar{n} = 1$. **a** Single (antibunched) photon source (Fock state), **b** laser (coherent) light, **c** single-mode thermal source (bunched)

consecutive photons for light sources with different photon statistics: coherent (laser) light, thermal light (bunching), and single (antibunched) photons.

Figure 4.2 shows probability histograms for different photon number n when the mean photon number $\bar{n} = 1$ for various types of photon sources (antibunched, coherent (laser) and thermal). For photon antibunching the second-order coherence function [3, 4, 28]

$$g^{(2)}(t) = \frac{\langle I(t)I(t+\tau) \rangle}{\langle I(\tau) \rangle^2} \quad (4.1)$$

should have a minimum at interphoton time $t = 0$. Here $I(t)$ is intensity, t and τ are two moments of time, and brackets denote time averaging.

For a single-mode field we may write [4]:

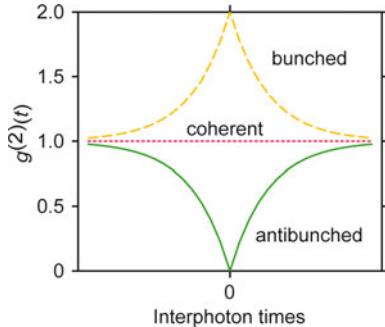
$$g^{(2)}(0) = 1 + \frac{\sigma^2 - \bar{n}}{\bar{n}^2}, \quad (4.2)$$

where σ^2 is the variance of the photon number distribution. For thermal (bunched) light $\sigma^2 = \bar{n}^2 + \bar{n}$, and $g^{(2)}(0) = 2$. For coherent (laser) light $\sigma^2 = \bar{n}$ (Poisson distribution) leading to $g^{(2)}(0) = 1$. For the light source with $\sigma^2 < \bar{n}$, $g^{(2)}(0) < 1$, and for a Fock state having fixed photon number n ,

$$g^{(2)}(0) = 1 - 1/n, \quad (4.2')$$

so for sources with $n = 2$, $g^{(2)}(0) = 0.5$. Figure 4.3 illustrates dependences of $g^{(2)}(t)$ for sources with different photon statistics.

Fig. 4.3 Dependence of the second-order coherence function $g^{(2)}(t)$ on the interphoton time t for light sources with different photon statistics



4.1.2 First Antibunching Experiments

The first intensity correlation experiments were performed by Hanbury Brown and Twiss in 1956, who used the coincidences between two detectors and measured $\langle I(t)I(t+\tau) \rangle$, showing photon bunching from a thermal light source [29]. In 1963 Glauber, who introduced quantum-mechanical field correlation functions of different orders [28, 30], pointed out that photon correlation experiments offer the possibility of observing the quantum-mechanical antibunching effect [4, 5].

Photon antibunching in fluorescence was predicted by Ehrenberg and Rigler (1974) [31], Carmichael and Walls (1975–1976) [32], Kimble and Mandel (1976) [33], and Cohen-Tannoudji (1977) [34]. Photon antibunching was observed experimentally for the first time by Kimble, Dagenais and Mandel in 1977 in the resonance fluorescence of single sodium atoms in laser-excited, collimated sodium beam [35–37], see Fig. 4.4a. The second antibunching experiment was made by Leuchs, Rateike and Walther [4, 38], also with sodium beam, see Fig. 4.4b. In 1985 antibunching was observed in a water-based solution of a fluorescent dye by Kask, Piksarv and Mets [39], see also [40]. In 1986 and 1987, antibunching in the resonance fluorescence of single cooled atomic ions using laser light excitation was observed by Bergquist et al. [41] as well as by Diedrich and Walther [42].

Development of single-molecule fluorescence studies in solids started from cryogenic temperatures in 1989 by Moerner and Kador [43] and in 1990 by Orrit and Bernnard [44]. A demonstration in 1993 of room temperature single-molecule fluorescence by Betzig and Chichester using near-field excitation [45] made possible later antibunching measurements of dye-molecule fluorescence at room temperature [46–50]. In 2000, Lounis and Moerner [51] and Fleury et al. [52] used the high stability of terrylene dye in para-terphenyl crystal to create a room-temperature SPS. Notably, [51] reported the first SPS “on-demand” by using pulsed-laser excitation. This opened a way for the future single-photon generation at definite times (in the case of SPS efficiency close to 1). The photostability of dye molecules used for SPS is critical, since photoblinking and photobleaching degrade the SPS efficiency. The fluorescence stability of terrylene and some other dyes was further

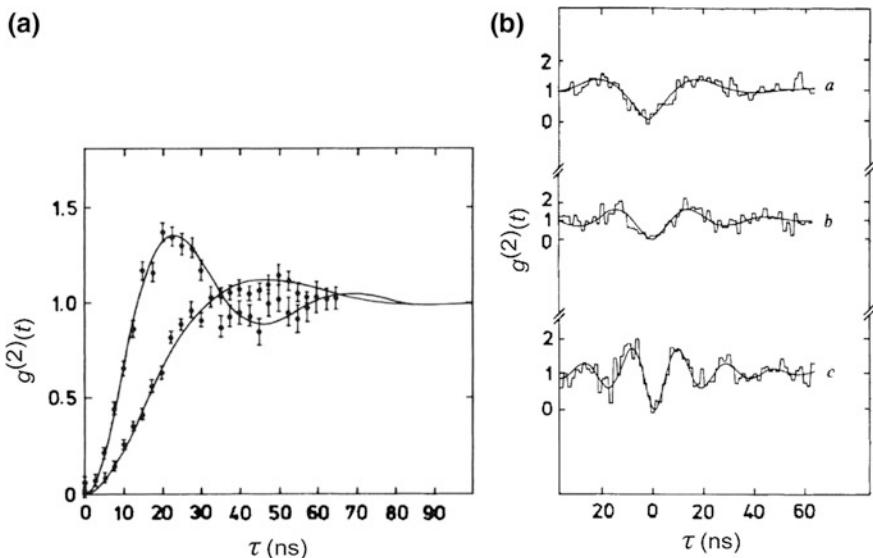


Fig. 4.4 Photon correlation measurements of fluorescence light from sodium compared with theory (solid lines) (reprinted from [4] with the permission of Nature Publishing group). The growth of $g(\tau)$ from $\tau = 0$ shows antibunching. **a** From experiments of Dagenais and Mandel [37], see also earlier [35, 36] of Kimble, Dagenais and Mandel. **b** Experimental points adapted from the results of Leuchs, Rateike and Walther, see also [38]

increased in polymethyl methacrylate [53, 54], polystyrene and polyvinyl chloride [54] films, under nitrogen [55], and in a liquid crystal host with oxygen depletion [56].

Other solid state systems for SPSs appeared around the year 2000, initiating an intensive research on SPSs worldwide. These materials included: nitrogen-vacancy (NV) centers in diamonds ([57, 58], see also [59, 60]); CdSe colloidal semiconductor nanocrystal quantum dots [61–63]; and epitaxial self-assembled semiconductor quantum dots [64, 65], see also [66–68].

4.1.3 Modern Trends on Development of Room-Temperature Single-Photon Sources

Several reviews provide good surveys of the field of single-photon sources (SPSs) [24, 69–85], see also books [86–88]. Furthest in development are SPSs based on heterostructured quantum dots [24, 70, 86, 87] placed into 2D-photonic crystal microcavities or into micropillars that permit excitation by electric field. In spite of their many advantages in studying cavity quantum electrodynamics, these SPSs are

very expensive: fabricated by molecular beam epitaxy and requiring cryogenic operating temperatures. Therefore, many researchers aim for *room-temperature* SPSs. The most popular single emitters for those are dye molecules, colloidal nanocrystal quantum dots and rods, various color-centers in diamonds and SiC nanoparticles. By embedding single emitters in photonic and plasmonic nanostructures, the following SPS performance parameters can be enhanced: spontaneous emission rate, directionality (coupling efficiency into a fiber system or external collection optics) and polarization selectivity (for efficient quantum qubit coding). Similar nano/microstructures can also be used for compact microlasers and on-chip integrated microlasers. The difference is in the concentration of the dopant: low dopant concentration for SPSs whereas microlasers require high concentration of the same emitters.

In spontaneous emission enhancement for SPS applications achieving two goals are equally important: (1) ultrafast spontaneous emission allowing SPS emission rate exceeding 10–100 GHz, and (2) enhancement of the single-photon collection efficiency, allowing total (detected) SPS count rates of Mcounts/s. For instance, using metamaterials can be a solution to reach the first goal, but to extract photons with high efficiency from metamaterials they should be combined with other nanostructures.

Physics of spontaneous emission enhancement is governed by the Purcell factor F_P [89, 90] which was initially introduced for microcavities and is defined as $F_P = \gamma_c/\gamma_0$, where γ_c and γ_0 are the emitter decay rates with and without a cavity. The Purcell factor can also be written as:

$$F_P = \frac{3Q(\lambda/n)^3}{4\pi^2 V_0} \left(\frac{|d| E(r)|}{|d||E(r)|} \right)^2, \quad (4.3)$$

where Q is the cavity quality factor, λ is the free-space wavelength of light, n is the cavity refractive index, V_0 is the cavity mode volume, d is the emitter dipole moment, and $E(r)$ is the local electric field at the position r of the emitter. Equation (4.3) predicts enhancement or diminishing of F_P , depending on the ratio $Q\lambda^3/(V_0n^3)$, and how the incident field is aligned with respect to the dipole moment, d .

Because high Q factors and mode volumes that are small compared to $(\lambda/n)^3$ are required to obtain $F_P \gg 1$, it is challenging to increase γ_c at optical frequencies. In recent years, however, semiconductor processing technology and other innovative procedures have enabled enhanced radiative decay rates in microcavities at visible and IR wavelengths, using both cryogenic and room temperature implementations. In addition to increasing γ_c , another advantage of placing the emitter inside a microcavity is the increased coupling efficiency into a single-mode fiber and the polarization selectivity in certain geometries of microcavities or nanoantennas. Even though the Purcell factor is the standard figure of merit for spontaneous emission rate enhancement, care must be taken when using (4.3) to describe emission rate enhancements for plasmonic resonances. In the case of plasmonic nanoantennas, the Purcell factor derived from quality factor and mode volume is

limited in describing emission changes due to plasmonic nanoantenna losses [91], although even with plasmonic nanoantennas it is used for evaluation of fluorescence decay rates of nanoemitters (see Sect. 4.3.3 for more details).

The most challenging aspect in SPS devices is creation of indistinguishable photons with transform limited linewidth single-photon emitters. Another challenging task is the detection of a single photon per each incident excitation pulse. Because of the low collection efficiency of modern SPSs, even “on-demand” SPSs based on pulsed laser excitation are not really “on-demand” [69]. See Sect. 4.3.1 for the definition of SPS efficiency.

4.1.4 Single-Photon Generation and Characterization Units

In a modern experimental implementation, single (antibunched) photons are produced by focusing a laser beam tightly into a sample area containing a very low concentration of emitters (less than 1 emitter/ μm^3). When this condition is satisfied, only one emitter becomes excited and will emit only one photon at a time, because of the finite times involved in absorption and fluorescence. Figure 4.5a shows a general schematic of an experimental setup for single photon generation, detection and characterization using a confocal fluorescence microscope. Figure 4.5b presents a desirable SPS device based on optical fibers (without a characterization unit).

A conventional single-photon generation unit consists of a confocal fluorescence microscope (Fig. 4.5a) with a high numerical aperture objective that focuses a pulsed or cw-laser beam on a sample with single emitters. This sample is mounted on a piezo-translation stage for raster scanning the sample through the focused laser beam. Fluorescence light is collected by the same objective. A dichroic mirror

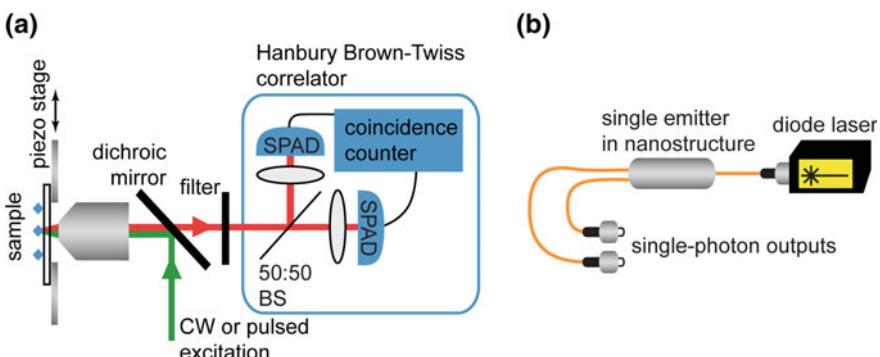


Fig. 4.5 **a** Schematic of a single-photon generation, detection and characterization unit (SPAD is single-photon counting avalanche photodiode module). A Hanbury Brown and Twiss correlator is used for photon antibunching measurements. **b** Schematics of a fiber-based single-photon source

reflects laser light and transmits fluorescence light. Interference filters with 6–9 orders of magnitude attenuation further reject the excitation laser light.

The single-photon detection and characterization unit (Fig. 4.5a) is a Hanbury Brown and Twiss intensity interferometer (correlator) [29], which consists of a nonpolarizing beamsplitter, two single-photon counting detectors (usually thermoelectrically cooled Si avalanche photodiode modules for visible and near-IR light), and start-stop electronics to measure time intervals between two consecutive photons in pairs to prove photon antibunching. A histogram can be built showing the number of occurrences $c(\tau)$ of photon pairs that are separated by time interval τ . To obtain the value of $g^{(2)}(\tau)$, which is proportional to coincidence count $c(\tau)$, normalization is required. The normalization can be deduced by calculation of $g^{(2)}(\tau) = c(\tau)/(I_1 I_2 \Delta t T)$, where I_1 and I_2 are the mean intensities on start and stop channels, Δt is the histogram time resolution, and T is the total acquisition time [63].

The characterization unit also contains electronics and software for confocal fluorescence imaging of single emitters. To increase contrast in confocal fluorescence imaging, either micrometer-size detector areas, additional apertures, or optical fibers serve as confocal microscope pinholes.

4.1.5 Overview of Next Sections

This chapter describes the state-of-the-art of room-temperature single-photon sources with photons exhibiting antibunching. Heralded single-photon sources [92, 93] obtained either in atomic two-photon radiative cascade [92] or with spontaneous parametric down conversion [93] will not be discussed here. A part of this review was presented by S. G. Lukishova as an invited paper at the SPIE conference “Emerging Technologies in Security and Defense II; and Quantum-Physics-based Information Security III”. Vol. 9254, paper 9254-05, <https://doi.org/10.1117/12.2066979> (October 2014, Amsterdam).

The structure of this chapter is as follows. Section 4.2 will outline the literature results on unbleachable single-emitter nanocrystals for room-temperature single-photon source applications, including nanocrystal quantum dots and rods (Sect. 4.2.1), diamond color centers (Sect. 4.2.2) and trivalent rare-earth ions (Sect. 4.2.3). New stable single-emitters for room-temperature SPSs will be outlined in Sect. 4.2.4. In Sect. 4.3, we will discuss the collection efficiency of a single-emitter on the dielectric interface (Sect. 4.3.1), and best microcavities (Sect. 4.3.2), nanoantennas (Sect. 4.3.3) and metamaterials (Sect. 4.3.4) for emitter fluorescence enhancement. Section 4.4 provides authors’ and other groups’ results on using liquid crystals to create single photons with definite circular and linear polarizations: photonic bandgap chiral microcavities in cholesteric liquid crystals (Sect. 4.4.1), emitter fluorescence enhancement with cholesteric microcavities

(Sect. 4.4.2), circular polarized fluorescence (Sect. 4.4.3), and linear polarization from single emitters in aligned liquid crystals (Sect. 4.4.4). Section 4.5 describes fiber-based SPSs. Section 4.6 concludes this chapter with comparison of different options and future directions for room-temperature SPS device implementation and use.

4.2 Unbleachable Single Emitters for Room-Temperature Single-Photon Sources

An ideal emitter for SPS application should avoid bleaching and blinking, should have a high photon counts rate and a narrow linewidth. For on demand SPSs the fluorescence lifetime should be less than the pumping period. In future applications, Fourier-transform limited, indistinguishable photons should be emitted. Electrically driven, on-demand excitation is preferable.

The most commonly used nanocrystals for room-temperature SPS applications are core-shell colloidal semiconductor nanocrystal quantum dots (NQDs) or rods, bulk diamond or nanodiamonds with different color centers, and nanocrystals doped with trivalent ions of rare-earths. For each of these emitters there have been demonstrations of long-operating times (photostability). It should be noted that electroluminescence was reported from NQDs [94–98] (including single NQD [94]), single NV center diamond [99, 100], and other single emitters [101, 102]. Thus there is no principal restriction for electrically driven excitation of these emitters, although in most papers on room-temperature SPSs simpler optical schemes are used.

4.2.1 “Giant” Colloidal Semiconductor Nanocrystal Quantum Dots and Dot-in-Rods

Colloidal semiconductor nanocrystal quantum dots (NQDs) of nanometer sizes (1–100 nm) [103, 104] consist of thousands of atoms (see Fig. 4.6a), but because of confinement of NQDs with the discrete, quantized energy levels of “a particle in the box” (Fig. 4.6b), they behave like single atoms. NQDs possess unique electronic and optical properties, for instance, their fluorescence wavelength is determined by their size: larger size NQDs emit at longer fluorescence wavelength (Fig. 4.6b).

Unlike epitaxially grown quantum dots, which only fluoresce single photons at cryogenic temperatures, NQDs are fabricated by relatively simple wet chemical techniques using precursor compounds dissolved in solutions. Most advantageously they fluoresce single photons at room temperatures although because of

inhomogeneous broadening the spectral width of NQD fluorescence exceeds 30–50 nm. Photon antibunching of NQD fluorescence was first observed in [61–63] and has since been reported by many groups. Typical NQDs are made of binary compounds such as II–VI (cadmium selenide (CdSe), cadmium sulfide (CdS)) or IV–VI (lead selenide (PbSe), lead sulfide (PbS)), etc., semiconductors. PbSe and PbS NQDs of the appropriate size fluoresce at optical communication wavelengths 1.3 and 1.55 μm [105–113]. Notably electroluminescence at telecom wavelengths was reported for PbSe NQD in a Si-microcavity [107].

The fluorescent properties of NQDs arise from the recombination of electron-hole pairs (excitons). However, excitons can also decay through a non-radiative process, reducing the fluorescence quantum yield. One of the methods used in improving efficiency and brightness of NQDs is adding shells of another, wider-bandgap semiconductor material around them. These NQDs are known as core-shell NQDs, for example, quantum dots with a CdSe core and ZnS in the shell. Core-shell NQDs may have quantum yield up to 100%. NQDs are available commercially.

The disadvantages of NQDs are (1) without special nanocrystal preparation, there can be occasionally strong, unpredictable blinking (fluorescence intermittency) and bleaching at the single NQD level; (2) clustering of NQDs in solutions; (3) multiexponential fluorescence decay of some NQDs; (4) fluorescence decay constants of NQDs (from the same solution, even on the same glass cover slips) can strongly vary from one dot to another. This prevents meaningful measurement of the Purcell factor of some NQD types. Although a precise mechanism has yet to be universally accepted, blinking is generally considered to arise from an NQD charging process in which an electron (or a hole) is temporarily lost to the surrounding environment (Auger ejection or charge tunneling) or captured to surface-related trap states. NQD emission turns “off” when the NQD is charged and turns “on” again when NQD charge neutrality is regained [114].

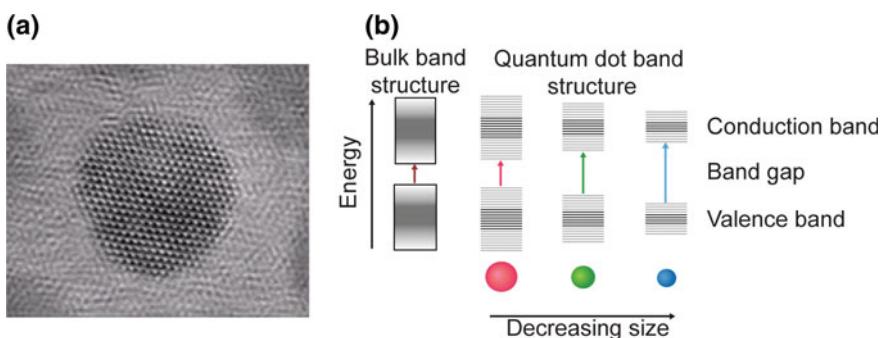


Fig. 4.6 **a** High resolution transmission electron microscope (TEM) image of CdSe–CdS core-shell quantum dot with suppressed blinking properties (9.2 nm total diameter with 4.4 nm core. Only Cd atoms are seen in this image). *Image credit* Ou Chen (permission from M. Bawendi). See also [115]. **b** Splitting energy levels in NQDs due to the quantum confinement effect, semiconductor band gap increases with decrease in size of the NQD

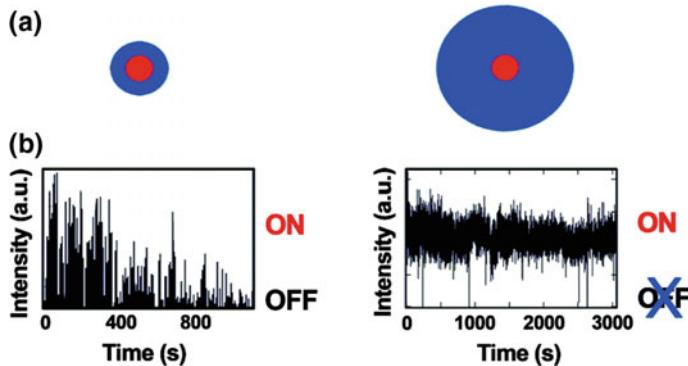


Fig. 4.7 (From [114] with permission). **a** Schematics of a conventional core-shell NQD and a “giant” NQD. **b** Time traces of a conventional NQD and a “giant” NQD

Recently, nonblinking and nonbleaching “giant” NQDs (g-NQDs) were discovered [114, 116–118] (see Fig. 4.7). In [114] the authors started with 3–4 nm NQD CdSe cores, and particles were grown up to a size of 15–20 nm by sequentially applying monolayers of inorganic shells. The g-NQDs were stable under continuous laser illumination (532 nm, 205 mW laser) at the single dot level. The photostability was measured by irradiating samples for several hours daily over a span of several days [114]. Photobleaching of the g-NQDs was not observed while half of the commercial NQDs (Qdot 655 ITK) photobleached within \sim 15 min. Figure 4.7 [114] shows time traces of a conventional, “blinking” NQD and a g-NQD with a very thick inorganic shell. Typically more than 70% of classically blinking NQDs have on-time fractions (fraction of total observation time that a single NQD is on) of <0.2 . In contrast, $>20\%$ of the g-NQDs are nonblinking —have an on-time fraction of >0.99 , and $>40\%$ of these g-NQDs have an on-time fraction of >0.8 . In addition, no clustering or aggregation of the g-NQDs was observed [114].

NQDs with a spherical core surrounded by a rod-like shell with a long length (similar to the size of g-NQDs) were fabricated in [119–122]. These quantum dot-in-rod (DR) nanocrystals (see Fig. 4.8) have stable fluorescence without blinking. These DRs aligned in a definite direction [122] can be used in SPSs with definite linear polarization (see Sect. 4.4.4).

4.2.2 Color-Center Diamonds

Because they possess long operating time (photostability) and short fluorescence lifetimes (of the order of 10 ns or less), diamond materials [123, 124] (both bulk and nanodiamonds) with different defects (color centers) are now considered the best candidates for practical, room-temperature single-photon source device

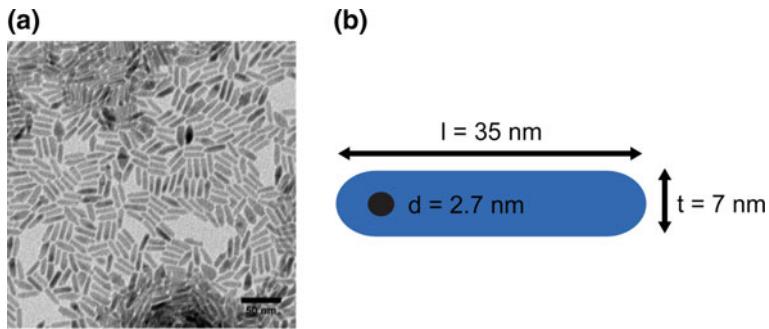


Fig. 4.8 (From [119] with permission). **a** Transmission electron microscope (TEM) image of the investigated DRs before dilution. **b** DR geometrical features determined from statistical analysis performed on TEM images. Shell length $l = 35 \text{ nm}$, shell thickness $t = 7 \text{ nm}$, and core diameter $d = 2.7 \text{ nm}$

implementation [57–60, 85, 125–127]. We note that color centers usually have lower fluorescence intensity in comparison with NQDs.

Defect-free diamond crystals are transparent within their wide bandgap (5.5 eV). The first experimental observation of individual defect center fluorescence in diamond at room temperature was reported in 1997 [128]. From more than 500 luminescent color centers in diamonds, only several have been identified as single-photon emitters with sufficient brightness. Among these bright emitters are the neutral and negatively charged nitrogen vacancy (NV^0 and NV^-), silicon vacancy (SiV), chromium related (Cr), nickel related (Ni), nickel-nitrogen (NE8), carbon (TR12) and others [127]. Table 4.1 shows a comparison of fluorescent properties of various color centers in diamond, see also [85, 127].

Color centers in diamonds can be produced by chemical vapor deposition (CVD) from a gas source (typically methane), by ion implantation, and by irradiation with high-energy electron beams and annealing at the temperatures greater than 800 °C. The smallest man-made nanodiamonds are produced either by detonation (5 nm) or via microplasma dissociation of ethanol vapor (2–5 nm) at near-ambient conditions.

From Table 4.1 (based on results from [57–59] and [127–150]) it is easy to see that nanodiamonds with Cr-related and some SiV color centers exhibited the best performance in terms of count rate, their lifetimes and linewidths (see also Fig. 4.9 for photoluminescence spectra from [127] and [151] of different color centers). These nanodiamond color centers are not yet available commercially, and there are relatively few groups producing these samples. However evaluation of NV^- and a typical SiV defect centers as SPSs for quantum key distribution using BB84 protocol over a short free-space transmission showed advantages of NV centers if SPS efficiencies are compared [152]. Usually SiV centers have low quantum efficiency and some of these centers are not stable. The use of nanodiamonds, rather than bulk single crystal, allows for a better collection efficiency of the photons. This efficiency enhancement arises because the subwavelength size of nanodiamonds. In the bulk

Table 4.1 Fluorescent properties of different color centers in diamond at the single center level (ZPL—zero phonon line wavelength, FWHM—emission bandwidth)

	ZPL (nm)	FWHM (nm)	Lifetime (ns)	Emission count rate (kcounts/s)	Quantum efficiency	References
TR12 in bulk diamond	470.5	~60 (measured at $T = 2$ K)	~3.6 (measured at $T = 2$ K)	—	—	[129]
NV ⁻ in bulk diamond	637	~100	11.6	6–8	0.7	[57, 58, 127, 128, 130]
NV ⁻ in nanodiamonds	637	~100	~25	50	For 50–150 nm size from 0.1 to 0.9	[59, 131–133]
H3 (N ₂ V)	503	~100	~24 ± 15	—	—	[134]
Unidentified centers	734	~4.1	~13.6	1800	—	[135]
	532	~70	~3	60	—	[136]
SiV in bulk diamond	738	~10	~1.1	1	0.05	[137]
SiV in nanodiamonds	738	~5–10	~1.1	6200 (on iridium substrate)	0.1	[138–140]
Cr in bulk diamond	749	~4	~1	500	0.29	[141]
Cr in nanodiamonds	749 and 756	~1.5 and 11	~1 and 14	3200 (756 nm)	>0.9 (756 nm)	[142, 143]
Ni/Si in bulk diamond	768	~5	~2	200	—	[144]
Ni/Si in bulk diamonds	773	~3	~1.1	77	—	[145]
NE8 in bulk diamond	802	~1.2	~11.5	75	0.7	[146]
	782	2	2	40	0.5 ± 0.2	[147]
	797	1.5	—	40	—	[148]
NE8 in nanodiamonds	793.7	~2	~2.1	35	—	[149]
GeV in bulk diamond	602	4–7	1.4	40, 120	—	[150]

case, the high refractive index of diamond reduces the light that can be collected from a color center.

It should be noted that most emitting diamond color centers exhibit a metastable or shelving state between the ground and excited states. Figure 4.10a shows a schematic representation of the energy level scheme of the NV center in diamond.

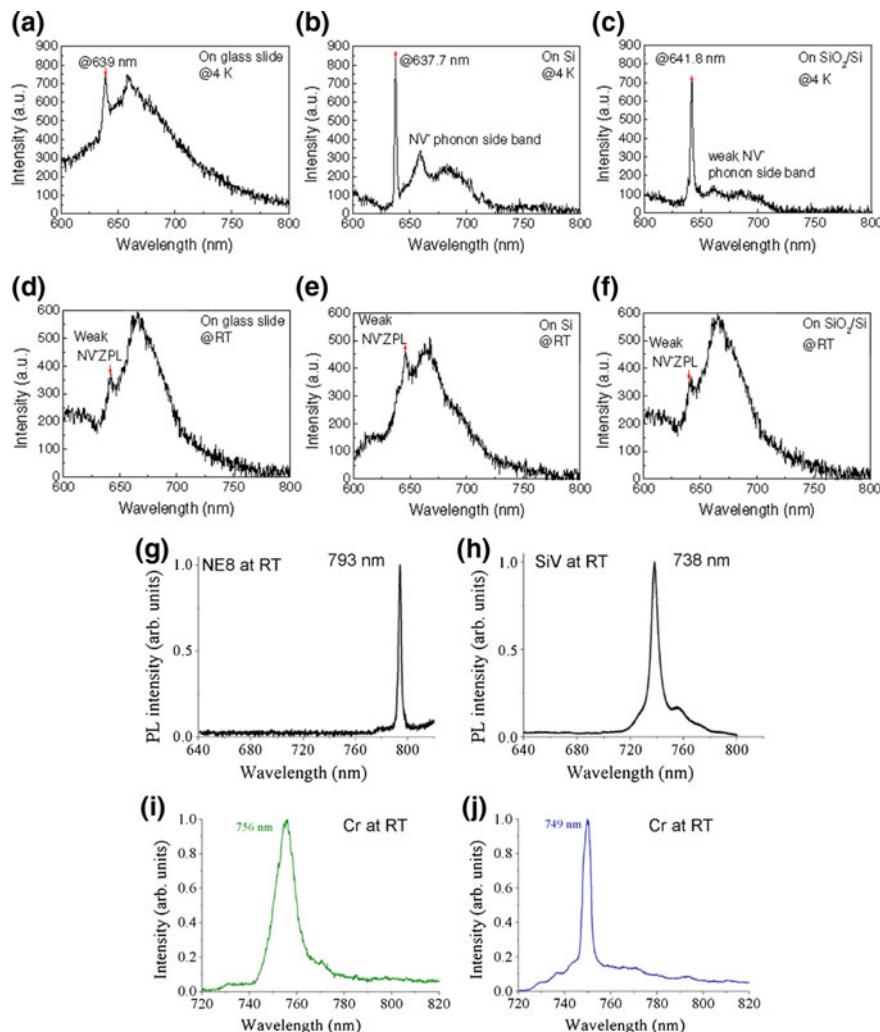


Fig. 4.9 Photoluminescence spectra of different color centers in diamond. Top 6 spectra of NV centers in nanodiamonds (a–f) [151] (reproduced with permission), measured at 4 K (a–c) and room temperature (d–f). The substrates were: glass (a, d), Si (b, e), and SiO_2 ($\sim 2 \mu\text{m}$)/Si ($250 \mu\text{m}$) (c, f). ZPL is shown for a NV^- -center. Bottom 4 spectra at room-temperature (by courtesy of I. Aharonovich [127]): the NE8 center (ZPL at 793 nm) (g), the Si-V center (ZPL at 738 nm) (h), chromium-related single-photon emitters with ZPLs of 756 and 749 nm, respectively (i and j)

Shelving on the metastable level ${}^1\text{A}$ is responsible for observed photon bunching characteristics with $g^{(2)}(\tau) > 1$ when the color center is excited above saturation. Figure 4.10b shows photon antibunching curves from a single NV center, below (blue curve) and above (yellow curve with bunching at higher interphoton times) saturation [153]. It should be noted, that Cr-related emitters can be either three-level

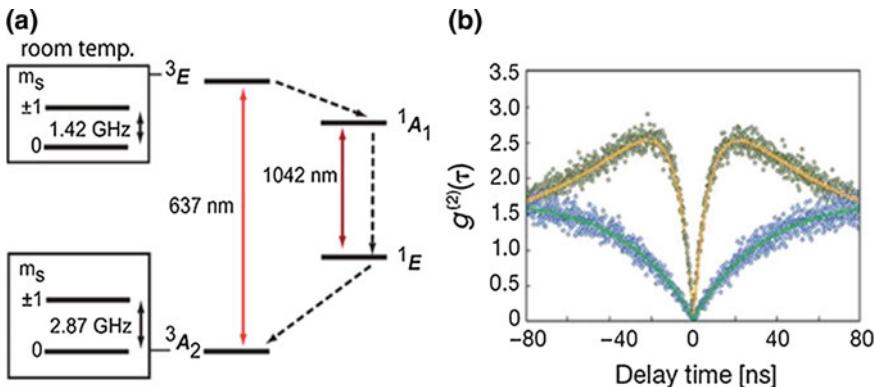


Fig. 4.10 **a** Schematic representation of the energy level scheme of the NV color center in diamond. **b** (Reproduced with permission from [153]). Example of $g^{(2)}(\tau)$ recorded below (blue lower curve) and above (yellow upper curve) saturation from a single NV center. Bunching is observed when the system is pumped above saturation

emitters with FWHM ~ 4 nm (Fig. 4.9j) or two-level emitters with FWHM ~ 11 nm (Fig. 4.9i). See also [127]. The two-level emitter has a quantum efficiency four times higher than that of three-level emitter, and does not exhibit bunching when pumped above saturation intensity. Polarization measurements showed that Cr-related and NE8 centers possess a well-defined linear dipole, in contrast to NV centers which do not emit fully polarized light at room temperature [127].

Only NV center nanodiamonds with variable number of color centers and with sizes starting from 10 nm and higher are available on the market. The characteristics of the first commercial SPS are reported in [153]: single-photon rates as high as 900,000 counts/s and with $g^{(2)}(0) < 5 \times 10^{-4}$. It is based on a NV nanodiamond grown on the facet of an optical fiber [154]. There has been another proposal to couple nanodiamonds to fiber by micromanipulation of nanodiamond onto fiber cores by an atomic force microscope (AFM) tip [155, 156].

Reference [157] reported the size reduction and effects on NV centers in nanodiamonds by air oxidation at elevated temperatures. The smallest nanodiamonds in which photon antibunching was observed were reported in [158] using “molecular sized” (< 2 nm diameter) SiV nanodiamonds obtained from meteorite powder. Luminescent NV center detonation nanodiamonds with sizes less than 6 nm have been observed in [159]. The influence of the surface increases with decreasing nanodiamond size. Nanodiamond fluorescence blinking was reported, e.g., for NV centers of 5-nm size crystals [160] and for SiV centers [139]. The blinking of the NV centers has been associated with charges trapped on the surface of nanodiamonds [86]; surface modification may inhibit blinking [159]. It should be noted that color centers in diamonds can be bleached as well.

By air oxidation at elevated temperatures of NV center nanodiamonds [157], the average height reduction of an individual crystal was 10 ± 1 nm/h at 600°C ,

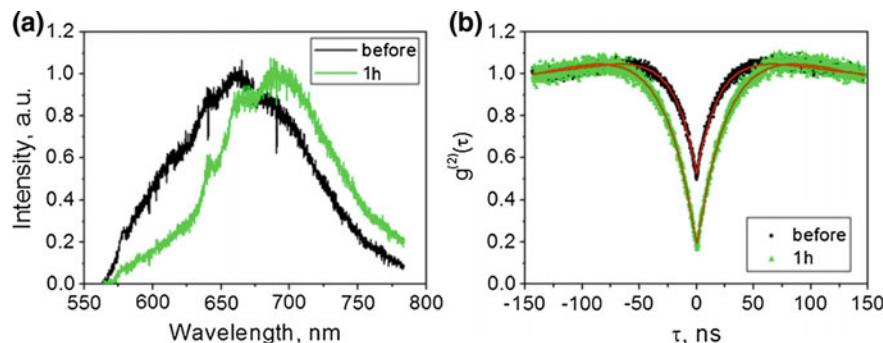


Fig. 4.11 (Reproduced with permission from [157]). **a** Normalized spectra of an NV center following a 1 h oxidation step. The black line is the spectrum without air oxidation, and the green one shows the spectrum after 1 h air oxidation. **b** Shows the corresponding $g^{(2)}$ curves. Note that the contrast of the antibunching feature increases, because of a decrease in background luminescence

4 ± 1 nm/h at 550°C , and less than 1 nm/h at 500°C . The oxidation process removed non-diamond carbon and organic material from the nanocrystal surface, which also resulted in decreased background fluorescence. The smallest nanodiamond hosting a stable NV center after size reduction was 8 nm. The authors used nanodiamonds from Microdiamant AG with an average size of 80 nm. Figure 4.11 shows changes of the spectrum (a) and antibunching curve (b) after an oxidation [157]. See also [161].

4.2.3 Nanocrystals Doped with Trivalent Rare-Earth Ions

In recent years, preparation and investigation of optical properties of nanocrystals with rare-earth [lanthanoid (lanthanide) according to the IUPAC nomenclature] ions attracted considerable interest in both fundamental research and practical applications, mostly for biomedical research, where such nanocrystals are used as fluorescence markers. Today, rare-earth-doped nanocrystals with sizes down to 10 nm are commercially available. These are comprised of trivalent rare-earth ion (TR^{3+}) impurities for upconversion luminescence through a combination of a TR^{3+} sensitizer (e.g. Yb, Nd, Er, or Sm) that initially absorbs the electromagnetic radiation and a second TR^{3+} activator (e.g. Er, Ho, Pr, Tm) ion in an optical passive crystal lattice that serves as the emitting element. Commercially available are also high-purity materials doped with rare-earths, including high-purity YAG crystals containing single-rare-earth ions.

TR^{3+} have a much wider fluorescence spectral range for operation (including telecom wavelengths 1.3 and $1.55\text{ }\mu\text{m}$) than is characteristic for color center diamonds. They are also more photostable than color centers. In addition to biomedical

applications (e.g., [162]), intensive research is focused on nanophotonics of upconversion luminescence in nanocrystals with sensitized energy transfer of TR^{3+} . For example, [163–165] explore the common, *forbidden* transitions with low oscillator strengths. Forbidden transitions of TR^{3+} can be used for single-photon source applications if microcavity or plasmonic resonance effects significantly diminish the lifetime of these transitions and increase the single-photon count rate. The long fluorescence lifetimes of common TR^{3+} forbidden transitions (the shortest is of the order of 10.6 μs [166]) are the main restriction in using TR^{3+} in SPS devices. Quantum communication systems need high bit rates of at least several tens of MHz and for this purpose fluorescence lifetimes should be in the nanosecond timespan.

Recently crystals including nanocrystals doped with *single* TR^{3+} ions were synthesized [167–177] owing to the *high purification of materials* as well as advances in synthesis methods, the control of nanocrystal size, composition, and crystal structure on the nanometer scale. Furthermore, photon antibunching under pulsed laser irradiation was reported recently at cryogenic temperature on a forbidden transitions $^3\text{H}_4 \leftrightarrow ^3\text{P}_0$ (488 nm excitation) and $^3\text{H}_4 \leftrightarrow ^1\text{D}_2$ (606 nm excitation) of Pr^{3+} ion in yttrium orthosilicate (YSO) crystallites (1–5 μm diameter) with a Pr doping level of 0.0001% [174]. The fluorescence lifetime curves showed a double-exponential fit in both cases with 2 and 162 μs fluorescence lifetime of the first transition and 25 and 277 μs of the second.

Antibunching was obtained at room temperature in much more intensive *allowed* transitions [167]. Figure 4.12a shows the energy levels of Pr^{3+} :YAG [178, 179], and its absorption and luminescence spectra are shown in Fig. 4.12b [180], see also [179]. Transition $4\text{f}5\text{d} \rightarrow ^3\text{H}_4$ is an allowed transition with a lifetime of ~ 18 ns at room temperature which can be reduced to 1 ns by temperature increase to 500 K [181]. Because strong absorption bands to $4\text{f}5\text{d}$ levels are located in the UV-region (see Fig. 4.12b), for this excitation, two visible-range photons are required (Fig. 4.12a). Emission from this level covers the spectral range 300–450 nm [167]. The quantum efficiency of this transition is close to unity even at room temperature [167]. Pr^{3+} was excited in [167] by a two-step upconversion excitation process involving an intermediate state (using $^3\text{H}_4 \leftrightarrow ^3\text{P}_{0,1,2}$ transitions). The first upconversion step uses a 453–488 nm laser (Fig. 4.12a). Once promoted into the $^3\text{P}_0$ state, the electron can be excited further into the $4\text{f}5\text{d}(2)$ band, nonradiatively decay onto the emitting $4\text{f}5\text{d}(1)$ level, and emit a photon (300–450 nm). The conditions under which photon antibunching from Pr^{3+} ions in YAG was reported in [167] were 488.2 nm laser excitation with 4.5 mW excitation power and an accumulation time of 1500 s.

There is another possibility to excite upconverted emission of Pr^{3+} in YAG in two steps. The first step is excitation of a metastable $^1\text{D}_2$ state with an orange laser (611 or 609 nm) followed by the second excitation step at 532 nm. This method was used by the same group as in [167] to demonstrate super-resolution microscopy of $\text{Pr}:\text{YAG}$ nanoparticles [178]. Even a *single color* optical pumping in the orange can be used though this scheme is much less efficient [167].

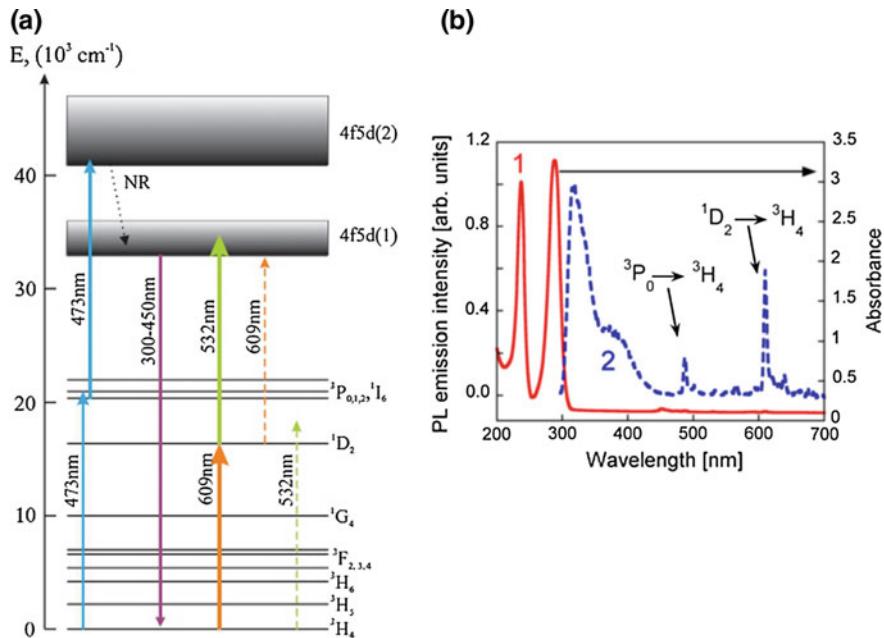


Fig. 4.12 **a** (Reproduced with permission from [178]): Energy level diagram of Pr^{3+} electronic states in YAG crystal. Two ways of stepwise excitation of upconverted UV fluorescence are shown. The excitation scheme is indicated by bold arrows. It is also indicated that the second excitation step taken with a 609 nm laser hits the very edge of absorption into the $4f5d(1)$ state and that a 532 nm laser cannot serve as the first excitation step (indicated by dashed arrows). **b** (Reproduced with permission from [180]): Absorption (curve 1) and photoluminescence spectra (curve 2) of Pr^{3+} :YAG with a 290 nm excitation wavelength. The micro-pulling-down-grown Pr :YAG sample with 0.1% Pr was used

Photon antibunching was also obtained by the same group of [167, 168] for the *allowed* transition of Ce^{3+} in YAG nanocrystals (Fig. 4.13) using pulsed excitation (second harmonic of a Ti: Sapphire laser) [177, 182]. The measured fluorescence decay time was 63.8 ns although at 700 K it can be reduced to 10 ns [181]. Figure 4.13a shows absorption (left curve) and fluorescence (right curve) spectra of the Ce^{3+} ion [183]. The energy level structure of Ce^{3+} [177, 183] is shown in Fig. 4.13b. Green-orange emission of Ce^{3+} can be obtained under blue or UV excitation. The quantum efficiency for these transitions is ~ 1 . The fluorescence lifetime for the same allowed transition of a single Ce^{3+} ion was also reported in [169], but no antibunching measurements were done in this work.

Pr^{3+} and Ce^{3+} ion implantation with high implantation yield was accomplished in [167, 177] that opens a feasibility of creation of nanopatterned rare-earth doping for on-chip photonic devices.

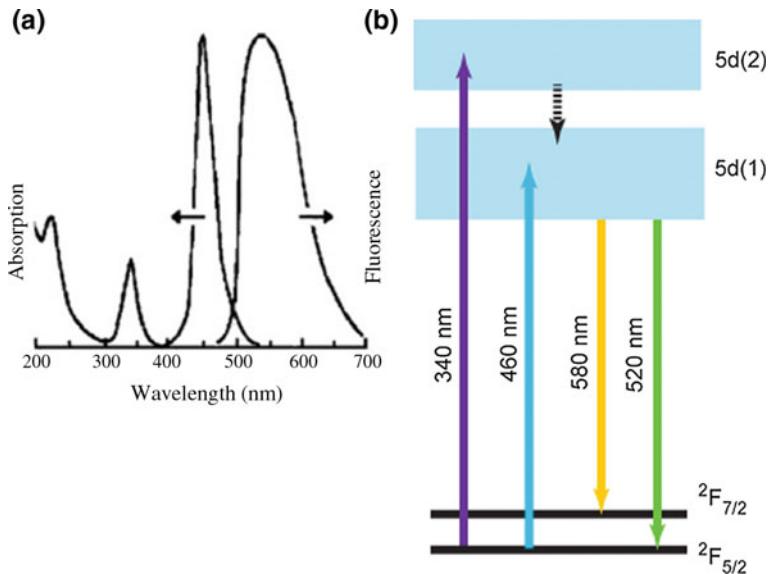


Fig. 4.13 **a** (Reproduced with permission from [183]). Absorption (left curve) and fluorescence spectra (right curve) of Ce^{3+} ion in YAG at 295 K. **b** Energy level structure of Ce^{3+} . Green-orange emission of Ce^{3+} can be obtained under blue or UV excitation

4.2.4 Other Types of Stable Single-Photon Emitters for Room-Temperature SPSs

Different types of materials with structure defects other than color centers in diamond have been proposed for room-temperature SPS applications. Most interesting among them are silicon carbide (SiC) [184–188], zinc oxide (ZnO) [189–192], carbon nanotubes [193], two-dimensional hexagonal boron nitride (hBN) [194], and borosilicate glass and quartz [195]. First, the number of defects in SiC appear promising. For instance, very bright fluorescence (up to 2×10^6 counts/s) and antibunching were observed from an intrinsic defect in SiC created by electron irradiation and annealing of ultrapure bulk SiC [184]. The fluorescence maximum was near 700 nm. For another defect in SiC with fluorescence in the near IR region (850–1200 nm) antibunching was also reported [185]. Electroluminescence of different SiC defects (450–650 nm and 850–1050 nm) was observed in [186]. Bright single-photon emitting diodes in the visible from SiC defects were fabricated (up to 3×10^5 counts/s) with photon antibunching observed from the electrically driven emitters both in continuous and pulsed modes at room temperature [187]. Room temperature single-photon emission from various defect centers in both bulk and nanostructured SiC was reviewed in [188].

Isolated defects in ZnO both in thin films and 50-nm-size nanoparticles with 660 and 725-nm fluorescence maxima have been used for single-photon generation

[189]. Single-photon emission from defect centers in 20-nm-size ZnO nanoparticles was also studied [190]. The emitters in [190] exhibit bright broadband fluorescence around 640 nm. Although the studied emitters showed continuous blinking, bleaching can be suppressed by a polymethyl methacrylate coating [190]. Stable electroluminescence from localized defects with maximum near 620 nm was observed for both ZnO films and nanoparticles in [192].

In [193], tunable room-temperature single-photon emission with single-photon purity 99% at telecom wavelengths (1300 and 1550 nm) from sp^3 defects in single-walled carbon nanotubes was reported. Earlier the same group obtained room-temperature single-photon generation in the 1100–1300 nm spectral range from solitary oxygen dopant states created by single-walled carbon nanotubes doped into a SiO₂ matrix.

Bright single-photon emission (up to 4×10^6 counts/s) at room temperature was observed from a defect in two-dimensional hexagonal hBN flakes with lattice structure similar to graphene [194]. The fluorescence maximum was 625 nm. The main advantages of 2D structures compared to bulk films and crystals are easier electroluminescence excitation and incorporation into optical cavities [196].

Antibunching was observed from defect centers in the borosilicate glass and quartz cover slips under 532 nm excitation [195]. It should be noted that in a microscope image, the emission from these defect centers is indistinguishable from spin-coated emitters ($\sim 10^4$ counts/s). The defects do not blink and have photoluminescence lifetimes of a few nanoseconds. Most centers photobleach within a minute while only a few remain emissive after several minutes. The emission spectrum is characterized by multiple peaks with a main peak at 583 nm for the glass and 600 nm for quartz, most likely due to coupling to a silica vibration [195].

A review on solid-state single-photon emitters was also recently published in [197]. See also [188].

4.3 Enhancement of Single-Photon Emission by Photonic and Plasmonic Structures

In this section we will discuss solutions for enhancement of both the single emitter: (1) spontaneous emission rate (diminishing fluorescence decay time) and (2) light extraction to the detector (photon count rate per second). See also [198] that provides a recent overview of different nanophotonic and nanoplasmonic structures that can be used for these purposes. In Sect. 4.4 we will present our experiments on using liquid crystal microcavities and single-emitter dipoles' alignment by liquid crystal structures for optimal excitation and deterministically polarized fluorescence of single emitters.

4.3.1 Collection Efficiency from a Single Emitter on a Cover Slip or Other Dielectric Surfaces

The probability P to one detect photon per excitation pulse depends both on the fluorescent properties of the single emitter and on the collection efficiency of the optical system. The value of P equals the product of $\eta\sigma\Phi C$ [199], where η is the fluorescence quantum yield, σ is the probability to excite the emitter at each pulse (σ is close to unity near saturation for nonblinking emitters), Φ is the detector efficiency, and C is the collection efficiency of the whole optical system. C depends on the collection efficiency of the objective, emission properties of the emitter at the glass interface, and optics (including interference filters) transmission. For an emitter at the air/glass interface, most of its fluorescence emission is directed into the medium with the high index of refraction (Fig. 4.14) [199]. The results of calculation in [199] for a two-dimensional emission dipole showed that the fraction of photons radiated into glass ranges between 82% and 88%, depending on the orientation of the dipole. For a microscope objective with a numerical aperture (NA) of 1.4, this means that approximately 72% of the fluorescence is collected within the collection angle α_{col} [199]. The value of C in a typical confocal microscope setup does not exceed 15%.

To increase collection efficiency of a microscope objective, solid immersion lenses (SIL) have been integrated into optical microscopes since 1990. Now they are commercially available and are used with single-emitters. E.g., count rates from stable NV centers in nanodiamonds were enhanced up to 853 kcounts/s [200] with a 0.9 NA objective using a 1 mm diameter solid immersion lens made of ZrO_2 ($n = 2.17$). This SIL was formed as a simple half-sphere with nanodiamonds spin coated to the flat surface with the curved surface directed to the optical objective. The concept of a SIL is very similar to traditional oil-immersion for microscope

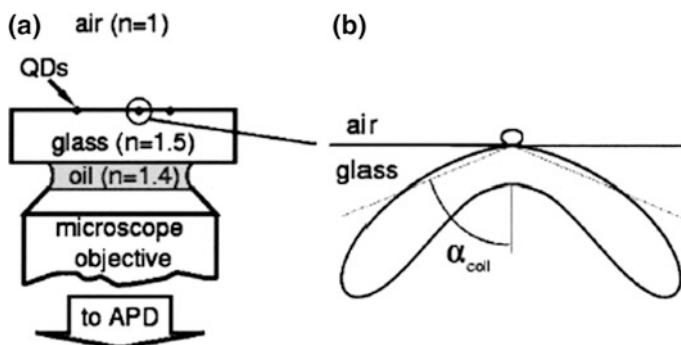


Fig. 4.14 (Reproduced with permission from [199]). Collection of the fluorescence of a single emitter. **a** Schematic overview for collecting the emitter fluorescence on glass by a microscope objective. APD is a single-photon counting avalanche photodiode. **b** Radiation pattern of a single emitter placed on a glass cover slip and covered by a 40 nm PMMA layer. Most of the power is radiated into the glass

objectives. Due to the high refractive index the NA is enhanced, providing larger collection angles and thus improved collection efficiency of the objective with the SIL. Calculation of the effect of the high refractive index at the SIL-air interface on the emission dissymmetry of a dipole located 10 nm away from a flat material surface with $n = 2.17$, showed that the dipole emits more than 86% of its emission into the direction of the SIL if oriented parallel to the surface and more than 75% if oriented perpendicular to the surface [200]. In [201], the fluorescence count rate from a NV center in bulk diamond was eightfold increased by using an 8 μm diameter SIL etched into the diamond.

In papers [202–204], a solution for almost complete collection of the photons from a single emitter, radiating into 4π solid angle is offered by using planar layered antennas. As was mentioned earlier, using only one dielectric interface leads to $\sim 12\text{--}18\%$ loss of into the upper half-space. Additional losses are incurred by emission into very large angles in the lower substrate, which are not accessible by the collection objective. In the experiments of [202], for two interfaces with refractive indexes $n_1 = 1.78$ (sapphire substrate) and $n_2 = 1.5$ (polyvinyl alcohol (PVA)) with terrylene dye molecules inside the PVA layer, $96\% \pm 3\%$ collection of the emitted light was reported using a microscope objective with NA of 1.65 ($\alpha_{\text{col}} \sim 68.8^\circ$). Even 99% collection efficiency for arbitrary orientation of the emitter dipole can be obtained by adding a metal mirror on the far side of the n_3 medium (air) (see Fig. 4.15a) to redirect and capture any leakage in the direction normal to the antenna plane [203, 204].

Figure 4.15b shows a proof-of-principle experiment of [204] with NQDs embedded between layers with the same refractive index $n_2 = 1.5$ (polymethyl methacrylate (PMMA) and PVA) on a sapphire cover glass. The third layer was air with $n_3 = 1$. In this demonstration, the gold mirror was placed at the cleaved end of a thinned optical fiber, so that the position of the mirror could be varied. The NA of the microscope objective was 1.65.

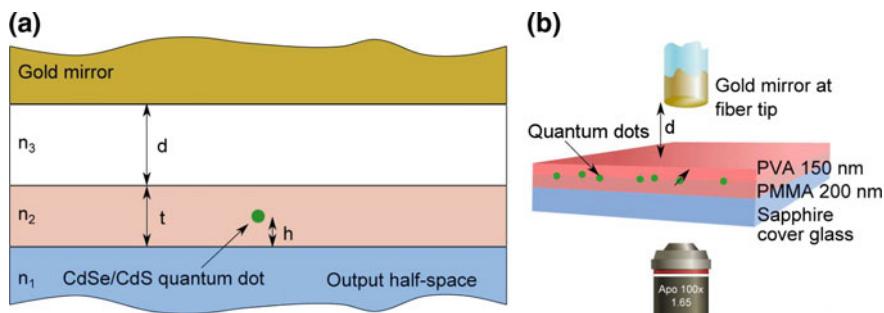


Fig. 4.15 (Reproduced with permission from [204]). **a** Layer arrangement of the metallo-dielectric antenna. **b** Proof-of-principle experimental realization of such an antenna with embedded CdSe/CdS giant quantum dots between two polymer layers (PMMA and PVA) of the same index of refraction, and a gold-coated end of a fiber is placed at variable separations from the top. A laser beam illuminates the sample through a 100 \times microscope objective

4.3.2 Single Emitters Within Microcavities Including Photonic Crystals

A cavity provides for tailoring of the spontaneous emission rate and radiative decay, modification of the spatial radiation pattern to increase the directionality and collection efficiency, narrowing of spectral lines, and definite polarization in certain microcavity geometries [205–211]. Two qualitatively different regimes of the time evolution in single-emitter fluorescence inside a cavity must be distinguished: the weak coupling regime and the strong-coupling regime [205–207]. In the weak coupling regime, an excited emitter loses its energy exponentially in time with a decay constant defined by the Purcell factor (4.3). The decay is enhanced when the emitter's transition frequency is in resonance with the cavity mode, or it may be strongly suppressed, when the emitter is off resonance. For the strong-coupling regime, energy is exchanged between the emitter and the cavity mode in an oscillatory manner at the Rabi frequency, where the oscillation is damped by the cavity decay constant (the spontaneous emission rate of the emitter may still be enhanced following a Purcell-like dependency). Different types of microcavities are reviewed in [211].

For efficient coupling to the cavity modes, the emitter luminescence linewidth must be sufficiently narrow. Therefore, the highest values of spontaneous-emission-rate enhancement by the cavities have been reported at cryogenic temperatures, although theoretically modelled benchmarks have not been reached yet. The most important issues towards that goal are increasing the Q factor, reducing the mode volume of micro(nano)-cavities, and placing a single-emitter with high precision exactly in the place of maximum electric field within the very small mode volume of the cavity.

2D-Microcavities for Epitaxial Quantum Dots, Color-Center Diamonds, and NQDs

The most compelling experimental result on spontaneous-emission-rate enhancement by a solid-state cavity was reported in [212] for the weak-coupling regime. Enhancement of photoluminescence peak intensity by up to a factor of ~ 700 proved the resonant coupling between epitaxial InAs/GaAs QDs and a 2D-photonic crystal (PC) nanocavity in a GaAs membrane at a temperature of 39 K.

Although the strong coupling regime is not ideal for SPS applications, in InAs the QD exciton lifetime was reduced by a factor of 145 in that regime at cryogenic temperature [213]. This strongly coupled cavity–QD system demonstrated strong quantum correlations in the form of photon antibunching. An epitaxial QD was grown inside a high Q ($\sim 30,000$) 2D PC GaAs nanocavity.

Different types of 2D PC micro/nano-cavities and their Q -factors are discussed in [214, 215]. These micro/nano-cavities from Si, GaAs or SiN can be prepared using electron-beam lithography and reactive ion etching. Figure 4.16a shows a micrograph of a typical Si 2D-PC with L3 defect nanocavity [216, 217]. Operating at telecom wavelength 1529 nm, this particular (L3) nanocavity consists of a triangular lattice of air holes on a silicon membrane, with a line of three holes omitted.

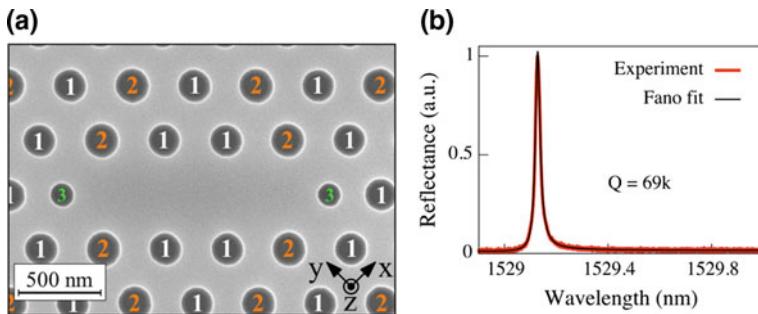


Fig. 4.16 (Reproduced with permission from [217]). **a** Scanning-electron micrograph of an L3 PC cavity optimized for vertical out-coupling. The holes are labeled according to their design radius. In this example, $r_1 = 89$ nm, $r_2 = 93$ nm, $r_3 = 62$ nm. Additionally, the holes labeled “3” are shifted outwards by 62 nm. **b** Measured resonance of a typical L3 PC cavity, displaying a Q of 69,000

The cavity’s terminating holes were reduced in size and slightly displaced to increase the cavity Q factor to 70,000. Additionally, the radii of selected holes surrounding the cavity were slightly increased to modify the cavity’s far-field emission pattern, and facilitate coupling in the out-of-plane direction.

Figure 4.16b depicts a spectral resonance in a typical L3 PC nanocavity at 1529 nm wavelength [217]. The nanocavity’s rectangular geometry also enhances polarization selectivity. Although the experimentally reached Q values of 2D PC nanocavities (Si-based) exceed one million [215, 218] numerical calculations [219, 220] predict even for low Q ($\sim 100\text{--}1000$) a Purcell factor in the range of 35–50. It can be further increased up to 75 by diminishing the microcavity mode volume using the inclusion of small high-index cylinders at the cavity [220].

Different types of microcavities (2D PC, 1D PC, microring, grating) have been fabricated with epitaxial quantum dots [221], in color-center bulk diamond [222–234] and bulk crystals doped with rare-earth ions [235–237]. Most measurements of the Purcell factor in these structures were also carried out at cryogenic temperature. For instance, the spontaneous emission into the zero phonon line (ZPL) of a NV center within a 2D PC was enhanced (as estimated by [223]) by a factor of 70, which corresponds to 70% of all photons being emitted in the ZPL and into the cavity mode (the NV fluorescence lifetime in these experiments diminished by a factor of 3 in comparison with a bulk diamond). In bulk diamond, the NV center ZPL contains only 3% of total fluorescence even at cryogenic temperatures (see also Fig. 4.9a, b). The Q value of a typical 2D PC in diamond was 3,000.

In [229], a 2D PC cavity was fabricated around a preselected, single, silicon-vacancy color center in diamond. The room-temperature cavity coupling gave rise to a resonant Purcell enhancement of the zero-phonon transition by a factor of 19, coming along with a 2.5-fold reduction of the emitter’s lifetime (Q factor was 430).

At room temperature, a nanodiamond with a single NV center was placed directly (by a nanomanipulation technique) on the surface of a gallium phosphide

(GaP) L3 2D PC cavity. A Purcell-enhancement of the fluorescence emission at the ZPL by a factor of 12.1 was obtained [238] relative to the ZPL peak intensities outside the cavity. The Q value was ~ 600 . In [239], NV center fluorescence emission enhancement was reported with estimated Purcell enhancement factor ~ 25 . A 50 nm size nanodiamond was placed inside one of the air holes surrounding the S1 2D PC cavity ($Q \sim 3800$) with the mode maximum localized in these holes. Another method of coupling NVs to microcavities is to position the microcavity over the desired emitter [240]. GaP PC membranes were lifted from their native substrate using a polymer stamp, and tip micromanipulation was used to position microcavities over a preselected NV centers. Spontaneous emission enhancement of two degenerate cavity modes coupled to the NV emission were measured as 2.2 and 7.0. The measured Q of freestanding cavities was as high as 6,000; the typical value was below 1,000 [240].

Room temperature emission from ensembles of PbSe and PbS NQDs in Si 2D PC microcavities was achieved in the work of [109–113, 241–243]. For instance, a more than tenfold photoluminescence enhancement near 1550 nm due to the cavity-nanocrystal coupling was observed in [112]. NQDs were attached to the designated area within a microcavity by employing a special surface treatment involving AFM-lithography and after that placing the sample into a NQD solution. The 2D PC microcavity resonances with PbSe and PbS NQDs were also observed in [109–113, 241–243]. The cavities' Q factors were ~ 800 –30,000. No fluorescence decay measurements were reported in these experiments.

1-D Bragg-Reflector Microcavities with Epitaxial Quantum Dots, Color-Center Diamonds and Rare-Earth Ions

For a one-dimensional microcavity in the strong-coupling regime at cryogenic temperature an “effective” Purcell factor of 61 was reported [221]. Photon antibunching was obtained as well. In this experiment, 1–4 μm -diameter planar micropillar Bragg reflectors were separated by a one λ thick GaAs microcavity with an InGaAs quantum dot layer grown in the middle of the cavity (the maximum Q factor reported in this work was $\sim 20,000$).

In the one-dimensional case (PC nanobeam [224, 225], see Fig. 4.17) an enhancement of the NV center ZPL fluorescence by a factor of ~ 7 was observed in [224], also by low-temperature measurements. The value of $Q = 6,000$ was

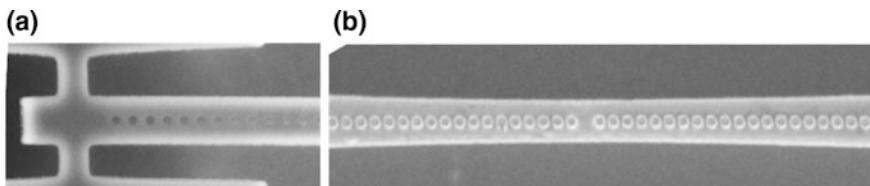


Fig. 4.17 (Reproduced with permission from [224]). SEM top view of a suspended photonic crystal nanobeam in a monocrystal diamond **(a)** as well as of the outcoupling region for transmission measurements **(b)**

reported for these experiments. The NV center diamond PC nanocavity of [226], attained by an advanced fabrication scheme, yielded a Purcell factor of 62 (strong Purcell regime) for the zero phonon line. The Q was 9,900 and the lifetime was reduced by the factor of 2 at cryogenic temperature. In [227] a circular bullseye diamond grating enabled high collection efficiencies of a single NV center.

Coupling of an ensemble of either Nd³⁺ [236] or Er³⁺ [237] rare-earth ions to the 1-D-triangular nanobeam-PC cavities fabricated in doped yttrium orthosilicate (YSO) host crystals was accomplished at cryogenic temperature. In [236], the observed reduction in lifetime (by a factor of 2.9) corresponded to an effective Purcell enhancement of 42. The difference between the observed lifetime reduction and the effective Purcell enhancement is due to the branching ratio between the available Nd³⁺ decay pathways, only one of which is enhanced by the cavity. The cavity Q was 4,400. Similar work with Er^{3+:}YSO 1-D-triangular nanobeam-PC cavities [237] yielded a lifetime reduction factor of 6, an effective Purcell factor of 53, and a cavity Q (for a resonance peak at 1536 nm) of 11,400.

1-D-Distributed Bragg-Reflector Microcavities with NQDs

Fluorescence from NQDs embedded inside 1-D, distributed Bragg-reflector (DBR) defect-layer microcavities at room temperature was investigated by several groups [78, 108, 216, 244–255]. The DBRs were fabricated from alternating quarter wavelength thick layers with different refractive indices [TiO₂–SiO₂, poly-vinylcarbazole (PVK), poly-acrylic acid (PAA), SiO₂–Si₃N₄]. See, e.g., Fig. 4.18a, b.

Enhanced spontaneous emission from CdSe/ZnS NQDs (610–620 nm emission wavelength) residing in DBR microcavities was reported for a NQD ensemble in the form of increased integrated fluorescence count rates (2.7-fold enhancement by the cavity [244]), and cavity-mediated fluorescence-decay-time reduction (e.g., 2.4 times in [251], 2.6 times in [252], and 2.7 times in [245]). In [252] the narrowing of the NQD ensemble photoluminescence linewidth down to a FWHM of 0.73 nm (from 30 nm) was reported at room temperature (see Fig. 4.18c). Q was ~835.

Photon antibunching from NQDs in DBR microcavities was observed in [78, 251–254] and in our work [216, 247, 248]. NQD patterning into micropillar arrays by means of electron beam lithography [256, 257] has been accomplished in [251, 252]. High Q pillar resonators with embedded colloidal CdSe/ZnS quantum dots or rods were fabricated from DBR structures by focused ion beam milling [253, 254]. Pillar microcavities reduce the losses into lateral waveguide modes. The Q factors measured from NQD photoluminescence reached values of 1,000–1,500 [253].

Yet another Bragg-reflector microcavity type with reduced waveguide modes is a chiral liquid crystal (cholesteric) microcavity that will be considered in Sect. 4.4.1. In such a structure we observed circularly polarized microcavity resonance and photon antibunching in the fluorescence from a CdSeTe NQD. There the cavity causes fivefold intensity enhancement at the NQD fluorescence maximum in a *circularly polarized mode of definite handedness* [258].

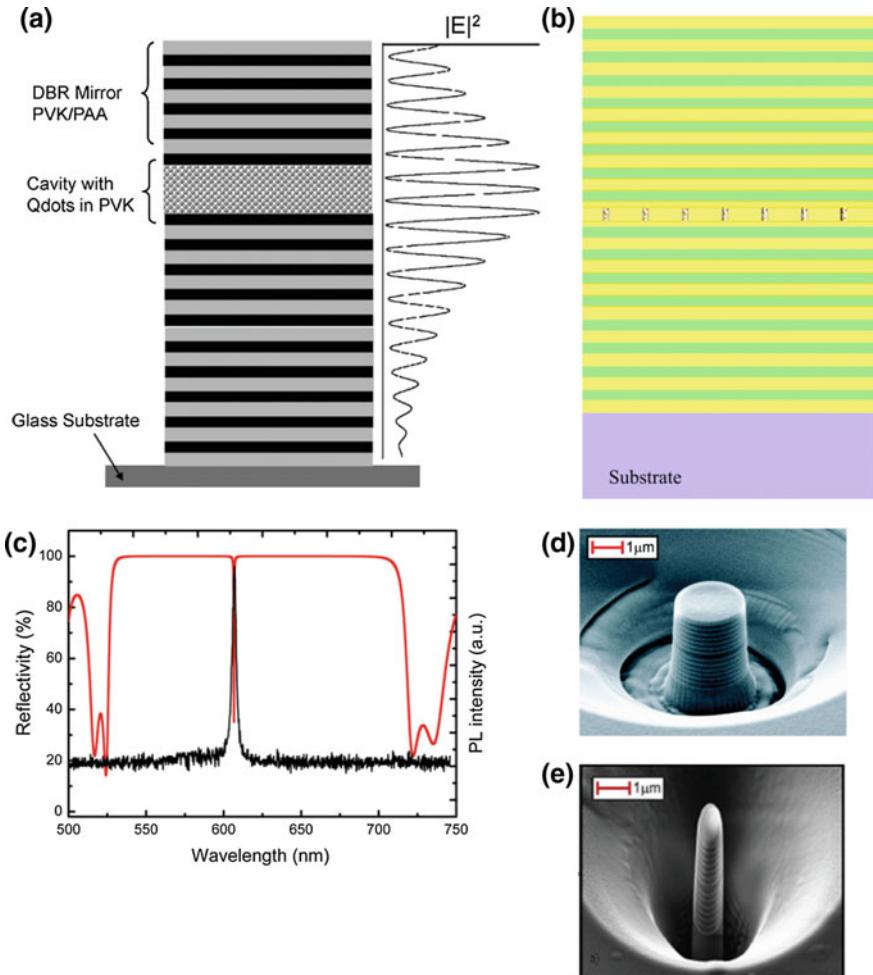


Fig. 4.18 **a** (Reproduced with permission from [245]). Schematic drawing of the all-polymer DBR microcavity structure (PVK/PAA) consisting of a half wavelength thick NQD layer with CdSe/ZnS NQDs sandwiched between two DBR mirrors. **b** (Reproduced with permission from [252]). Schematic drawing of the $\lambda/2$ planar cavity consisting of two $\text{SiO}_2\text{-TiO}_2$ Bragg mirrors and of CdSe/ZnS NQDs localized in an array of pillars in the middle of the cavity. **c** (Reproduced with permission from [252]). Overlap of the reflectivity and the photoluminescence spectra collected from the top of the implemented device **(b)** at room temperature. **d, e** (Reproduced with permission from [253]). SEM images of micropillar resonators with round (top) and elliptical (bottom) cross sections

Fabry-Perot and Fiber-Optics Based Microcavities with Nanodiamonds

A small-mode-volume, fiber-optics-based Fabry-Perot microcavity with a NV center diamond was experimentally realized in [259–261]. Its schematic [261] and cavity emission spectrum are presented in Fig. 4.19a, b. The concave mirror

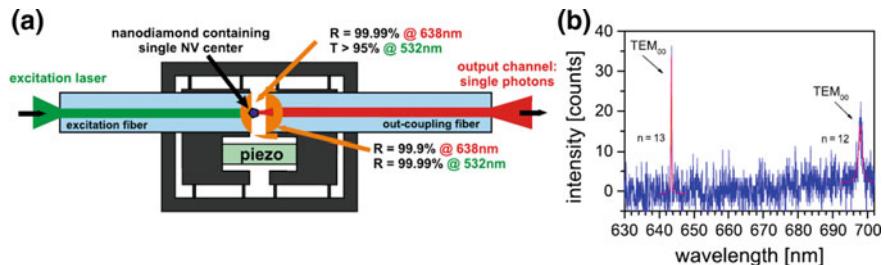


Fig. 4.19 (Reproduced with permission from [261]). **a** Schematic of a fiber-based Fabry-Perot microcavity consisting of two concave mirrors fabricated on the facets of optical fibers. **b** Cavity emission spectrum of cavity with the smallest effective length $4.1\ \mu\text{m}$ with the given concave profiles of the two fibers

separation was controlled via piezo-electric crystal, so the emission wavelength is readily tunable. The output mirror prepared by focused-ion-beam milling of a fiber facet had a parabolic shape, was $1.2\ \mu\text{m}$ deep, and had a radius of curvature of $14.1\ \mu\text{m}$. The input mirror with a nanodiamond containing a single NV center was $1.9\ \mu\text{m}$ deep and had a radius of curvature of $71.6\ \mu\text{m}$. By coupling this emitter to this cavity, an increase of the spectral photon rate density by two orders of magnitude was achieved compared to the free-space emission of the NV center. For this cavity length the linewidth was $10.1\ \text{GHz}$, and the spectral photon rate density was $420\ \text{photons}/(\text{s GHz})$ in contrast to the free-space emission of $4.1\ \text{photons}/(\text{s GHz})$. The cavity mode volume was $4.1\ \mu\text{m}^3$. Photon antibunching was also recorded in these experiments. A similar cavity with NV nanodiamonds ($1.24\ \mu\text{m}^3$ mode volume and $Q \sim 3,000$) was investigated at $77\ \text{K}$ [262]. The signal from individual zero-phonon-line transitions was enhanced by a factor of 6.25 and the overall emission rate of the NV centers was increased by 40% compared with that measured from the same center in the absence of the cavity.

Whispering Gallery Mode Resonators with Nanodiamonds and NQDs

Extensive research has been carried out on coupling single-emitter fluorescence with whispering gallery mode (WGM) resonators with high Q values (microspheres, microdiscs, microtoroids, etc.), see reviews [263–268]. Adiabatically tapered optical fibers can provide perfect coupling (99.99%) with these structures. Reference [269] reports on the coupling of a single CdSe/ZnS NQD chemically bound to the surface of a single, glass, WGM microsphere ($3\text{--}10\ \mu\text{m}$ in diameter) with the Purcell factor ~ 5 at $10\ \text{K}$ temperature by comparing decay times ($Q \sim 7,434$).

At room temperature, multiple microcavity resonances in NQD fluorescence were observed. Similarly, multiple resonances were reported for NV center nanodiamonds coupled to WGM polystyrene and silica microspheres [270–272]

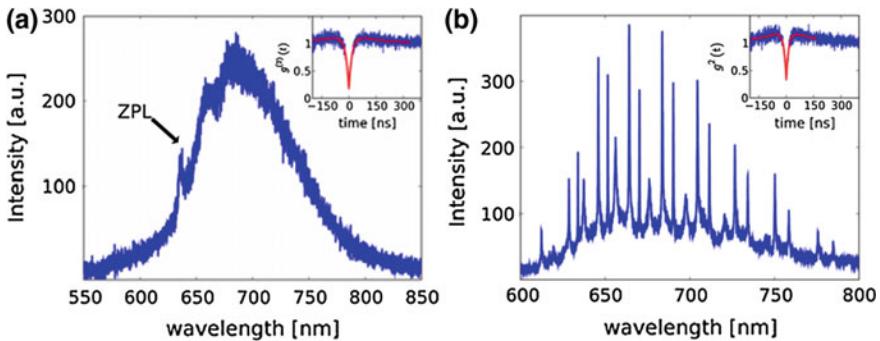


Fig. 4.20 (Reproduced with permission from [270]). The fluorescence spectra from a single NV center in nanodiamond: **a** on the coverslip; **b** coupled to a precharacterized microsphere, the peaks correspond to high Q WGMs. The insets show the corresponding second-order correlation functions with the data fitted with a three level model

(see Fig. 4.20) and microtoroids [272, 273]. An interesting approach to WGM microstructures is fabrication by two-photon direct laser-writing on a photoresist doped with single-emitter nanocrystals [272, 274], see Fig. 4.21.

Using this method, quantum photonic circuits were built [272, 274], which consist of quantum emitters, resonators, and waveguides (Fig. 4.21). The functionality of this concept has been proven by showing single photon emission from a NV center inside a WGM resonator by measuring the statistics of the photons at the output ports of a waveguide [272, 274].

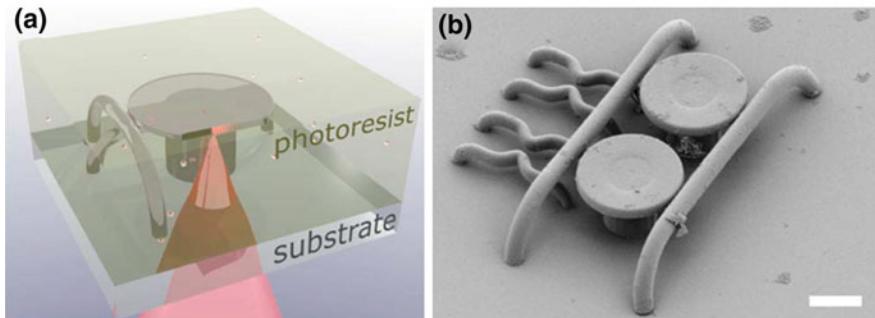


Fig. 4.21 (Reproduced with permission from [274]). Direct laser writing in nanodiamond photoresist. **a** Sketch of the direct laser writing process. A femtosecond laser beam is focused into the photoresist in order to polymerize well defined 3D structures. **b** Scanning electron micrograph of such a structure after development containing waveguides, couplers and microdisc resonators. Scale bar is 5 μm

4.3.3 Single Emitters Within Plasmonic Nanoantennas

Another way to enhance single-emitter transition rate is offered by metal plasmonic nanoantennas [275–291]. At present, plasmonic nanoantennas provide both the highest reported emission rate enhancements and the highest photon flux increase from single emitters [284–288]. The optical properties of the metal structures are strongly affected by surface plasmon polaritons, i.e. electromagnetic excitations propagating at the interface between a dielectric and a metal. They arise from the coupling of electromagnetic fields to oscillations of the metal’s electron plasma. A prominent example with a pronounced impact on single-molecule and single nanoparticle detection is metal surface enhanced Raman scattering: there enhancement factors of the order of 10^{14} – 10^{15} have been observed.

Plasmonic nanoantennas serve to spatially enhance and localize fields, and modify the excitation rate and the radiative decay rate when placed close to single emitters. But they can also cause undesirable losses, leading to an increase in the non-radiative decay rates of these emitters. This interplay of rates can lead to a strong modification of the emission characteristics of emitters. In plasmonic structures, the Purcell factor combines contributions from an increased radiative decay rate and from an increased non-radiative decay rate due to metal losses [91, 285]. It is therefore critical to specify the fraction of energy emitted as radiation, known as the radiative quantum efficiency. From knowledge of the Purcell factor and the radiative quantum efficiency, the enhancement in the radiative decay rate can be obtained $\Gamma_{\text{rad}}^{\text{a}}/\Gamma_{\text{rad}}^0$, where $\Gamma_{\text{rad}}^{\text{a}}$ and Γ_{rad}^0 are the radiative decay rates with antenna and in a free space.

Metal plasmonic nanoantennas of different geometry are fabricated by electron-beam lithography from different materials. Figure 4.22 shows different types of such nanoantennas [281], see also [90, 280]. Nanocrystals containing single-emitters are placed within nanoantennas by either spin coating or atomic-force microscope (AFM) tip manipulation [165, 292–299]. Figure 4.23 shows our success in manipulating nanodiamonds [216, 247]. The three images at the left [(a)–(c)] show diamond nanocrystals (identified as 1 and 2) moved by an AFM tip. The right image (Fig. 4.23d) shows creation a “smiling face” from an initially chaotic distribution of nanodiamonds using an AFM tip.

Bowtie Plasmonic Nanoantennas

Bowtie nanoantenna consists of two triangle tips separated by a nanometer-sized gap, see Fig. 4.22. In [284] dye molecules were spin coated onto a sample with bowtie nanoantennas. In our paper [247], we spin coated CdSeTe NQDs (Qdot 800 ITK organic, Invitrogen) with fluorescence maximum at 790 nm onto a sample with gold bowtie nanoantenna arrays. Some NQDs appeared in the gap of bowtie nanoantennas. We were able to record (Fig. 4.24a) photon antibunching from a NQD within a 30-nm-gap bowtie nanoantenna (see insert) in spite of polycrystalline gold photoluminescence (Fig. 4.24b). NQD fluorescence within a bowtie is much

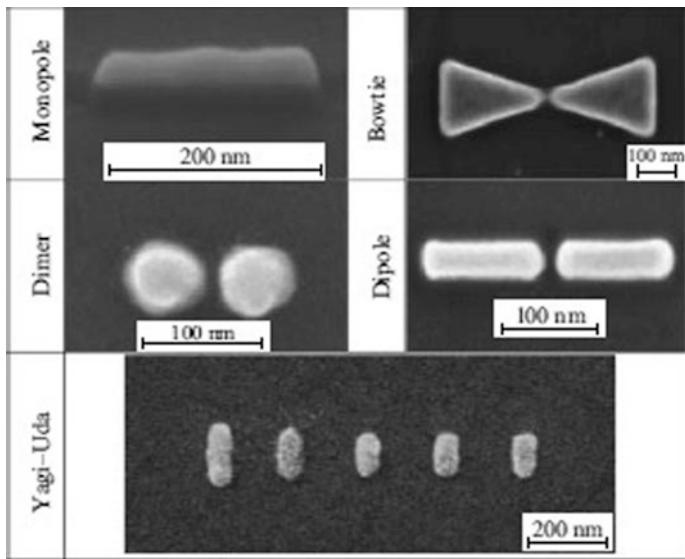


Fig. 4.22 (Reproduced with permission from [281]). SEM images of optical nanoantennas

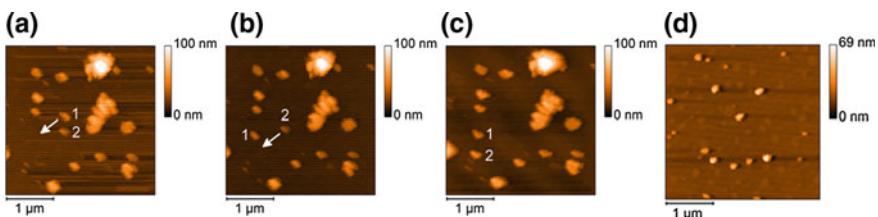


Fig. 4.23 (Reproduced from [216]). **a–c** (three images) and **d** (one image): Results of AFM-tip manipulation by nanodiamonds (see explanation in the text)

brighter than of gold that can be seen from time traces of single NQD fluorescence (Fig. 4.24c) [247].

Photoluminescence from gold exhibits a broad spectrum in the visible and near-infrared overlapping with plasmon resonances as a result of thermally deposited, nanometer-scale rough polycrystalline gold [297, 300–302]. The polycrystalline gold photoluminescence originates from the intra-band light absorption in gold and the radiative decay enhancement in the presence of nanometer-scale roughness. Recent papers on monocrystalline gold flakes report suppression of this gold photoluminescence [303, 304].

The fluorescence rate Γ_f from a single emitter is determined by the excitation rate Γ_{exc} and the intrinsic quantum yield Q_i of the emitter. Q_i is defined as $\Gamma_{\text{rad}}^0 / (\Gamma_{\text{rad}}^0 + \Gamma_{\text{nr}}^0)$, where Γ_{rad}^0 and Γ_{nr}^0 are the intrinsic radiative and non-radiative decay rates. In the absence of the antenna and far from fluorescence saturation, the

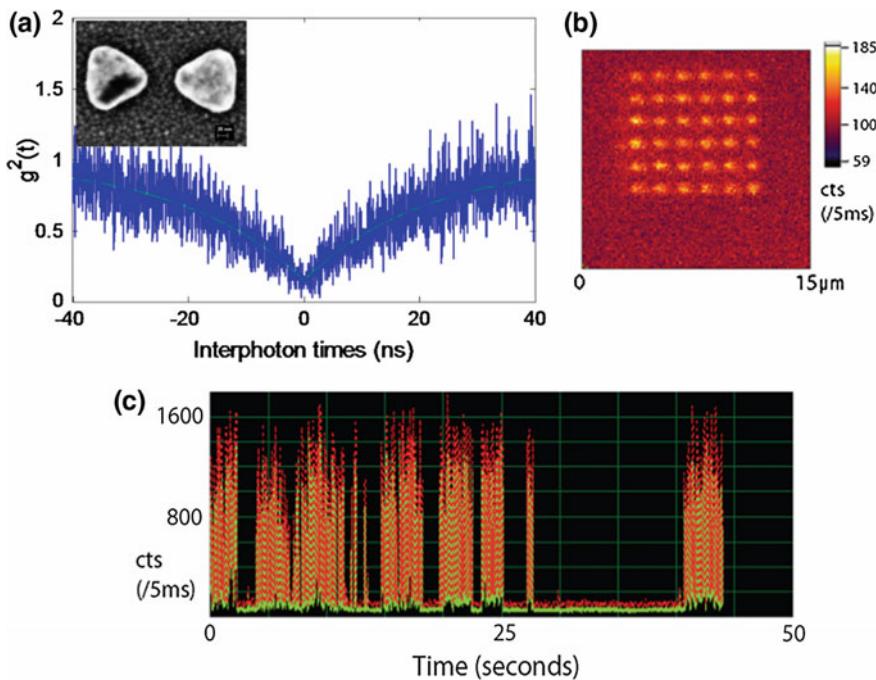


Fig. 4.24 (Reproduced from [247]). **a** Photon antibunching from a CdSe/CdTe NQD (790 nm fluorescence maximum, 633 nm excitation wavelength) within a 30 nm gap gold bowtie nanoantenna (insert shows SEM image of a typical bowtie nanoantenna). **b** Gold photoluminescence of a typical nanoantenna array at 633 nm excitation. **c** Time traces of NQD photoluminescence within a gold bowtie nanoantenna showing NQD blinking

fluorescence rate is given by $\Gamma_{\text{fl}}^0 = \Gamma_{\text{exc}}^0 Q_i = \Gamma_{\text{exc}}^0 \Gamma_{\text{rad}}^0 / (\Gamma_{\text{rad}}^0 + \Gamma_{\text{nr}}^0)$. The antenna typically leads to an increase in Γ_{exc} to Γ_{exc}^a due to local field enhancement. An increase in Γ_{exc} is accompanied by an increase of Γ_{rad} to Γ_{rad}^a . However, the presence of the antenna also increases Γ_{nr} to Γ_{nr}^a because of non-radiative energy transfer from the excited emitter to the nanoantenna. The key to having efficient antennas for fluorescence enhancement is to minimize non-radiative losses in the metal (quenching). Simulations and experiments by the Novotny group [278, 289] showed that for short distances between the emitter and a metal antenna, the decrease of quantum yield wins over the increase of the excitation rate thereby quenching the fluorescence of the emitter. Both the experiment and the theory showed on the example of single dye molecules and 80 nm gold nanosphere that the fluorescence enhancement reaches a maximum at a distance of $z = 5$ nm. For shorter distances fluorescence is quenched. The authors of [290, 291] showed by numerical modeling that, using different nanoantenna structures manufacturable with today's nanotechnology, it is possible to increase the radiative decay rate by three orders of magnitude while keeping a quantum efficiency larger than 80% in the near-infrared regime. They examined the competition between the radiative and

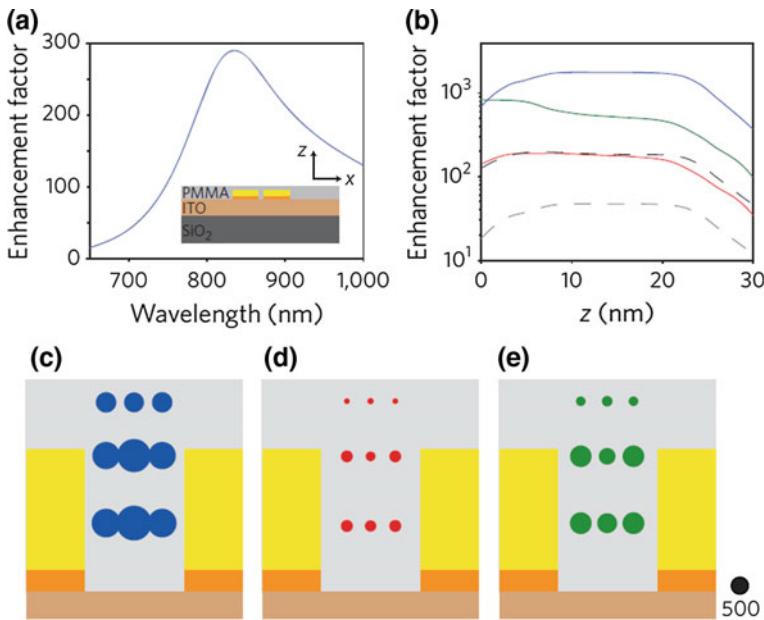


Fig. 4.25 (Reproduced with permission from [284]). **a** Spectrum of calculated electric field intensity enhancement versus wavelength in the centre of a bowtie with a 14-nm gap. *Inset* the simulated structure (side view) consists of a SiO₂ (refractive index $n = 1.47$) substrate, a 50-nm layer of indium tin oxide (ITO) ($n = 2$), and a 30-nm layer of PMMA ($n = 1.49$). The gold bowtie structure is 20 nm thick on a 4-nm layer of titanium. **b** Radiative (red, lowest solid curve) and non-radiative (green, middle solid curve) enhancement factors along the center of the gap at a wavelength of 820 nm; z is a measure of the distance above the ITO/PMMA interface. The black top dashed line represents the local optical intensity enhancement at 780 nm. The blue curve (top solid curve) shows the fluorescence enhancement factor for molecules with a quantum efficiency of 2.5% and the grey dashed line that for molecules with a quantum efficiency of 100%. **c–e** Illustration of the simulated structure (side view, section through the two triangle tips) showing regions of fluorescence [(c), blue], radiative [(d), red] and non-radiative [(e), green] enhancement factors for a molecule emitting at a wavelength of 820 nm

nonradiative processes in the presence of the antennas as a function of wavelength and antenna geometry, and compare radiative decay enhancement and antenna efficiency at different wavelengths for gold, aluminum and copper nanoantennas.

For gold bowtie nanoantennas numerical modeling of radiative and nonradiative enhancement factors $\Gamma_{\text{rad}}^{\text{a}}/\Gamma_{\text{rad}}^0$ and $\Gamma_{\text{nr}}^{\text{a}}/\Gamma_{\text{nr}}^0$ was performed by the Moerner group [284]. These results are presented in Fig. 4.25. In Fig. 4.25b the top (blue) solid curve shows the fluorescence brightness enhancement factor of 1,690 for emitters with $Q_i = 2.5\%$ and the lowest dashed (grey) line shows an enhancement factor of 50 for emitters with $Q_i = 100\%$. In the experiment, a fluorescence brightness enhancement factor of 1,340 for dye molecules (820 nm maximum fluorescence wavelength, $Q_i = 2.5\%$) was obtained [284]. The difference in enhancement factors for different Q_i occurs because the total fluorescence enhancement ratio $\Gamma_{\text{fl}}^{\text{a}}/\Gamma_{\text{fl}}^0$ is equal to $f_Q \Gamma_{\text{exc}}^{\text{a}}/\Gamma_{\text{exc}}^0$, where f_Q is the enhancement of quantum efficiency by the

antenna. Calculations of [284] showed that for $Q_i = 2.5\%$, $f_Q = 9.3$. Radiative and nonradiative enhancement factors are presented by a lowest solid (red) curve and a middle (green) solid curve in Fig. 4.25b. The black top dashed line represents the local optical intensity enhancement at the dye absorption wavelength of 780 nm ($\Gamma_{exc}^a/\Gamma_{exc}^0 \sim 181$).

At the center of the bowtie gap the radiative factor $\Gamma_{rad}^a/\Gamma_{rad}^0$ was 187 and the non-radiative factor $\Gamma_{nr}^a/\Gamma_{nr}^0$ was 577. In the vicinity of the antenna, the non-radiative process due to metal loss thus dominates the radiative process, resulting in a quantum efficiency $\Gamma_{rad}^a/(\Gamma_{rad}^a + \Gamma_{nr}^a)$ of 25%.

For large enhancement of radiative and nonradiative processes decreasing single-emitter decay times should be observed. The fluorescence decay time decrease by a factor of >28 was observed experimentally in [284]. The time accuracy of a measuring system did not permit the measurement of decay times shorter than ~10 ps for the dye molecule fluorescence, so in reality the factor by which the decay time decreased could be much higher.

Bullseye Nanoantennas/Gratings

A bullseye nanoantenna or grating [227, 298, 299] consists of concentric slits or circles of metal or dielectric. References [298, 299] describe the method of fabricating a hybrid metal-dielectric bullseye nanoantenna with a NQD positioned in its center. The period of the circles was 600 nm and the slit width was 200 nm. This device efficiently directed photons from the NQD into a small divergence angle perpendicular to the nanoantenna surface. These measurements have shown that more than 20% of the photons are emitted into a small numerical aperture of NA = 0.25, and more than 35% of the emission is collected with a larger NA = 0.65, allowing large collection efficiency with very simple collection optics, or even directly into a fiber. Reference [299] also reported photon antibunching of the NQD within this nanoantenna.

Similar bullseye nanopatterning will be considered in the subsection of Sect. 4.3.4 treating metamaterials with embedded NQDs.

Plasmonic Nanogap and Nanopatch Antennas

To obtain even larger field enhancement, the nanoscale gap between a nanoantenna's metal arms should be diminished to sub-10 nm size, which is not easy to fabricate with current electron beam lithography. Promising geometries for overcoming these challenges are: (1) the gap-mode plasmonic nanocavity [305, 306] with emitters placed between a metal nanowire lying parallel to the substrate and a metal substrate; and (2) the plasmonic nanopatch antenna (NPA), which consists of emitters situated in a vertical dielectric gap between a metal nanoparticle of a given shape (cube, triangle, etc.) and a metal film [285–288, 307–313].

Reference [305] demonstrated radiative emission rate enhancement of the fluorescent organic dye molecule tris-(8-hydroxyquinoline) aluminium (Alq3) coupled to the nanoscale gap between a silver nanowire and a silver substrate. The gap spacing ranged from 5 to 25 nm, and the emission rate enhancement factors

relative to a bulk dielectric film (with a nominal decay rate of $1/46\text{ ns}^{-1}$) approached 1,000.

In [285–288, 307–312], solutions of gold and silver nanocubes with different sizes (50–100 nm) were synthesized by methods of colloidal chemistry [313]. The plasmonic resonance of such NPAs can be tuned from 500 to 900 nm by varying the size of the nanopatch [285].

In [287] the NPA consisted of colloidally synthesized silver nanocubes deposited over a 50 nm thick silver film. The cubes and film were separated by an ~ 5 nm self-assembled polyelectrolyte spacer layer, coated with a dilute layer of fluorophores sulfo-cy5 carboxylic acid (Cy5). By sweeping the plasmon resonance but keeping the field enhancements roughly fixed, fluorescence enhancement (comparison of spectral maxima) was demonstrated exceeding a factor of 30,000. Detector-limited enhancement of the Cy5 fluorophore spontaneous emission rate by a factor of 74 (in comparison with glass) and by a factor of 5 (in comparison with a silver film) was also demonstrated. These values were obtained for nanopatch resonances close to the excitation wavelength.

Further work of this group with a similar NPA [285], but with a dye with a long intrinsic fluorescence lifetime (~ 600 ns) and an intrinsic quantum efficiency of 20%, (ruthenium metal complex (Ru) dye) enabled fluorescence lifetime measurements on ensembles of dye molecules. Spontaneous emission rate enhancements exceeding 1,000 were obtained for NPA gaps ≤ 8 nm. For gaps $d < 7$ nm, quenching of the dye emission by the metal becomes a significant loss mechanism. Although in general it is difficult to distinguish radiative enhancement from metal quenching because experimentally only the total emission rate is accessible directly, the authors evaluations showed both high quantum efficiency $\Gamma^a_{\text{rad}}/(\Gamma^a_{\text{rad}} + \Gamma^a_{\text{nr}}) > 50\%$ for gaps ≥ 8 nm and directional emission (84% collection efficiency). In these experiments the dye fluorescence spectrum overlapped with the gold film-cube resonance, but the NPAs were excited non-resonantly. A fluorescence intensity enhancement factor above 60 was obtained experimentally for 8 nm gaps.

Using a 6 nm diameter CdSe/ZnS NQD within a NPA (silver nanocube coupled to a gold film) [286], an increase in the spontaneous emission rate by a factor of 880 and simultaneously a 2,300-fold enhancement in the total fluorescence intensity with a high radiative quantum efficiency of 50% were achieved, as well as directional emission. In these experiments, an ultrafast (<11 ps) yet efficient source of spontaneous emission, corresponding to an emission rate exceeding 90 GHz, was obtained. For these measurements, only NPAs with plasmon resonances around the NQD emission wavelength (625–635 nm) were selected. Furthermore, controlling the dimensions of the nanocubes and the gap thickness opens the possibility for resonances in the near infrared for spontaneous emission sources at telecommunication wavelengths. A reduction of the fluorescence lifetime to 12 ps (by a factor of 625) was reported in [312] for 10 nm size CdSe/CdS/ZnS NQDs in a cavity between aluminum and a triangular silver nanoprism NPA. A 3 nm thick aluminum oxide isolating layer was deposited on the aluminum.

It should be noted that in the most experiments of papers [284–287, 312], ensembles of emitters were involved. Further work is needed at the single-emitter level, leading to antibunching demonstration as in [288]. In our experiments with photon antibunching from single NQDs within bowtie nanoantennas as presented in Fig. 4.24, NQDs with 790 nm fluorescence maximum matched the gold-bowtie plasmonic resonance. Our experiments showed that for these particular NQDs the fluorescence decay rate varies significantly from one dot to another, even without nanoantennas, on glass substrates, so no reliable measurements of fluorescence enhancement by nanoantennas for these NQDs was possible. In contrast, NQDs with shorter wavelength fluorescence, e.g., near 600 nm, are more identical relative their fluorescence lifetimes, but usually only silver-based nanoantennas have plasmonic resonances in this spectral region, see, e.g. [286, 288]. Yet silver as a material is, in difference with gold, environmentally unstable.

4.3.4 Single Emitters with Metamaterials

Metamaterials (from the Greek word “meta-”, μετά-meaning “beyond”) are composite artificially engineered materials containing subwavelength features that control the macroscopic electromagnetic properties of the material. In particular, metamaterials with hyperbolic dispersion [314–332] (in which the real parts of the permittivity ϵ_{par} and ϵ_{perp} have different signs), have broadband “supersingularity” of the photonic density of states:

$$\omega^2/c^2 = k_{\text{par}}^2/\epsilon_{\text{perp}} + k_{\text{perp}}^2/\epsilon_{\text{par}}, \quad (4.4)$$

where subscripts “ par ” and “ perp ” indicate the directions parallel and perpendicular to the plane of anisotropy, respectively, and k is the wave vector.

As a result, such a medium allows the propagation of high- k modes with arbitrarily large wavevectors, which causes enhanced and highly directional spontaneous emission. This property of hyperbolic metamaterials (HMM) make them attractive for SPS application. In addition, spontaneous emission enhancement in metamaterials is broadband. The HMM is practically realized in two types of structures: alternating subwavelength-thick layers of metal and dielectric, or as an array of nanowires embedded into a dielectric host matrix [318].

HMM with Arrays of Nanowires in Dielectric Host

Figure 4.26a (from [317]) shows a HMM coated with dye-doped polymeric film. The material exhibited a hyperbolic dispersion with the effective values of permittivity $\epsilon_{\text{par}} = 4.99 + i0.22$ and $\epsilon_{\text{perp}} = -0.15 + i1.07$.

Reference [317] demonstrated a shortening of the fluorescence decay rate of dye molecules deposited on HMM in comparison with the same dye deposited on metal films. The maximum decay rate enhancement factor observed was ~ 6 , shown in

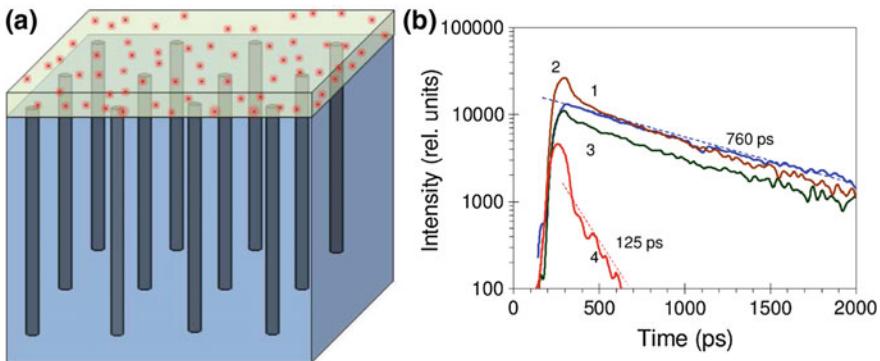


Fig. 4.26 (From [317]). **a** Schematic of the HMM with alumina membrane with circular silver nanowires coated with dye-doped polymeric film. **b** Emission kinetics in the dye IR140/PMMA films deposited on (1) the top of pure alumina membrane, (2) gold film on glass, (3) silver film on glass, and (4) silver-filled alumina membrane (HMM)

Fig. 4.26b. This experiment also showed that photon emission into the hemisphere above the HMM/dye film interface is strongly reduced. The high photonic density of states in a hyperbolic metamaterial gives rise to preferential and enhanced spontaneous emission into the HMM [317].

Nanopatterning of Multilayered HMM with a Slit Grating

To overcome the challenge of collecting emission from HMMs, [330–332] employed nanopatterned HMMs. In [330], a HMM was made from Ag and Si multilayers (each layer was 10 nm thick, the total thickness 305 nm). This HMM was patterned with a nanoslit grating (period 80–200 nm; width 40 nm), and dye molecules mixed in PMMA were spin-coated onto the grating surface to a thickness of 80 nm. The spontaneous emission rate of rhodamine dye molecules was enhanced 76-fold and the emission intensity of the dye increased by ~80-fold compared to a HMM without nanopatterning.

Nanopatterning of Multilayered HMM with a Bullseye Grating

In [331], a high-index contrast bullseye grating with a half period ~125 nm on the top of a layered HMM provided both enhancement in the spontaneous emission rate by a factor of ~10 (in comparison with glass) and efficient light extraction with a factor of ~20 enhancement (in comparison to a HMM without bullseye grating). NQDs were placed in the middle of the HMM. The HMM structure consisted of seven periods of alternating layers of aluminum oxide (~20 nm thickness) and silver (~12 nm thickness). 1–2 nm thick Ge seed layers were deposited on top of each aluminum oxide layer to wet the silver layer.

Nanopatterning of Multilayered HMM with a Photonic Hypercrystal

In another paper, the same group [332] used a two-dimensional photonic hypercrystal on top of a layered HMM to enhance the fluorescence decay rate and the light extraction from NQDs embedded in the HMM. The fluorescence decay rate was enhanced by a factor of 20 (in comparison with glass) and light extraction from the HMM was enhanced by a factor of ~ 100 in comparison with HMM without the photonic hypercrystal. Seven alternating layers of Ag (~ 9 nm) and Al_2O_3 (~ 16 nm) were deposited (Fig. 4.27a). The active layer of NQDs (635 nm center of emission) was embedded in the middle of the 5th dielectric layer by spin coating. As in [331], ~ 1 nm thick Ge seed layers were deposited before each Ag layer. A photonic hypercrystal (Fig. 4.27c) composed of an array of nanoholes was patterned in the HMM with focused ion beam milling. The holes were milled down to the NQDs layer.

Multilayered HMM Ceramics

In [324] a SPS based on NV color centers in nanodiamonds on the surface of a HMM was reported (see Fig. 4.28). The HMM was fabricated as an epitaxial metal/dielectric superlattice consisting of 10 pairs of layers of CMOS-compatible ceramics: titanium nitride (TiN , 8.5 nm thick) and aluminum scandium nitride ($\text{Al}_x\text{Sc}_{1-x}\text{N}$, 6.3 nm thick). This work demonstrated 4.7 times count rate enhancement in comparison with nanodiamonds on a glass slip. The average and largest decreases in fluorescence lifetime were a factor of 4 and 11.4, respectively. The photon count rate enhancement from this HMM was greater than predicted by modeling the NVs as dipoles above the HMM. The authors suggested that the extra enhancement could be explained by the presence of surface defects or scattering objects on the HMM surface that may result in outcoupling into the far-field of the modes propagating into the HMM.

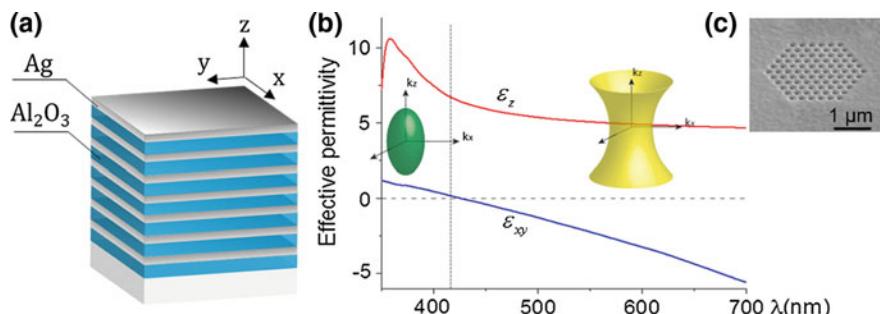


Fig. 4.27 (Provided by the authors of [332]). **a** Schematic of HMM composed of 7 periods of Al_2O_3 and Ag. **b** Effective permittivity tensor components. Dashed line at 426 nm marks the transition wavelength from elliptical dispersion to hyperbolic dispersion regime. **c** SEM micrograph of a photonic hypercrystal composed of an array of nano-holes patterned in the HMM with focused ion beam milling

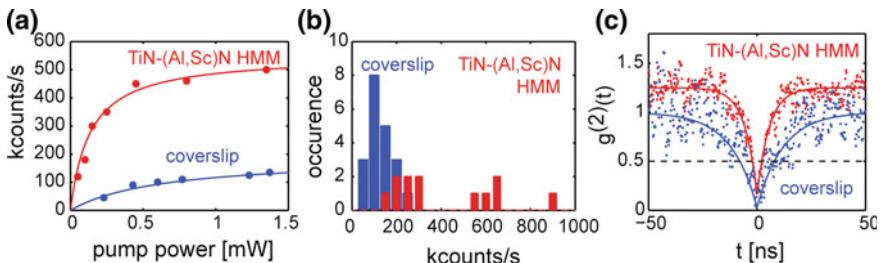


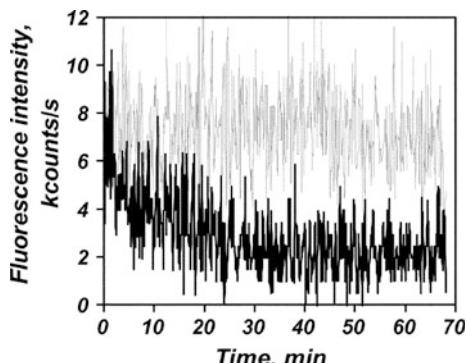
Fig. 4.28 (Reproduced with permission from [324]). Collected single-photon count rates (corrected for background emission) from NV centers in 50 nm nanodiamonds. **a** Typical saturation curves, **b** histograms of count rates for nanodiamonds on glass coverslip (blue) and HMM (red). Multiple maxima are evident. The average enhancements for the first and second statistical maxima are 1.8 and 4.7, respectively. **c** Second-order correlation function $g^{(2)}(t)$ of a representative nanodiamond with single NV center on top of glass coverslip (blue) and HMM (red)

4.4 Single Emitters in Liquid Crystals: Single Photon Sources with Definite Circular and Linear Polarizations

Liquid crystal (LC) materials (including glassy liquid-crystal oligomers, fluid-like monomers, and polymers) doped with single emitters have following properties that can be used for efficient SPSs with definite linear or circular polarization [56, 79, 108, 258, 333–342]:

- (1) Using LC technology with oxygen depletion can prevent bleaching of some emitters (Fig. 4.29) [56], which we describe below.
- (2) Planar-aligned cholesteric-liquid-crystal (CLC) hosts are one dimensional (1-D) photonic-band-gap microcavities [108, 343, 344] tunable by electric field or temperature [345]. We describe the general preparation of these microcavities in Sect. 4.4.1. The chiral structure of the CLC host material provides circular polarized output photons with definite handedness [108, 258, 333]. We describe

Fig. 4.29 (Reprinted with permission from [56]). Fluorescence bleaching behavior of an assembly of terrylene dye molecules as a function of time and in oxygen depleted LC host (upper grey curve) and in LC host without oxygen depletion (black lower curve)



our work with glassy oligomeric CLC microcavities in Sect. 4.4.2, and with monomeric CLC microcavities in Sect. 4.4.3.

- (3) Planar-aligned rod-like nematic LC molecules or defects in LCs can align the emitter molecular dipoles preferentially for maximum excitation efficiency. Definite direction molecular alignment will also provide output photons with definite linear polarization [122, 333, 337, 339, 341, 342]. Section 4.4.4 describes our work and that of others groups on SPSs with definite linear polarization based on nematic LCs doped with single emitters.
- (4) The LC photoalignment technique [346, 347] permits creation of a complicated alignment distribution on a small area, e.g., optical fiber facets or inside microstructured optical fibers.

Figure 4.29 shows our results of preventing terrylene dye bleaching in the oxygen-depleted LC host [56]. Over the course of more than one hour, no dye bleaching was observed in such a LC host (upper curve). Monomeric 5CB cyanobiphenyl LC was saturated, prior to cell assembly, with helium in a sealed glove-box for one hour. Oxygen that is mostly responsible for this dye bleaching was displaced by helium during this procedure. The lower curve shows dye fluorescence in an LC host without oxygen depletion. Both experiments used the same incident intensity on the dye molecules. The concentration of terrylene molecules in both cases was two orders-of-magnitude higher than in single-molecule experiments with photon antibunching.

4.4.1 Photonic Bandgap Chiral Microcavities in CLC

Planar-aligned chiral CLC structures exhibit a 1-D photonic bandgaps for the handedness of circularly-polarized light where the electric field vector follows the rotation of the CLC molecular director (Fig. 4.30a). The stop band is centered at wavelength $\lambda_c = p(n_e + n_o)/2$, where p is the pitch of the CLC spiral structure, and n_e and n_o are extraordinary and ordinary refractive indexes. The bandwidth of the transmission stop band is given by $\Delta\lambda \sim p(n_e - n_o)$ [348].

Lasing in such CLC structures occurs at a band edge of this stop band [343, 344, 349–354], because of the density of states is suppressed within the stop band and enhanced at its edge. In a similar way, to enhance single-emitter fluorescence in such a microcavity, the single-emitter fluorescence maximum should correspond to the band edge of the CLC structure stop band [258].

For development of CLC hosts which form a chiral photonic bandgap microcavities tuned to the single emitter fluorescence band, two main aspects are important: (1) properly choosing the concentration (or ratio) of different LC components and (2) providing planar alignment of the CLC. For the monomeric (fluid-like) mixtures, the stop band position λ_c of the photonic bandgap is defined roughly by $\lambda_c = n_{av}/(c_m HTP)$, where n_{av} is the average of the CLC ordinary and extraordinary refractive indices, c_m is the weight concentration of a chiral additive

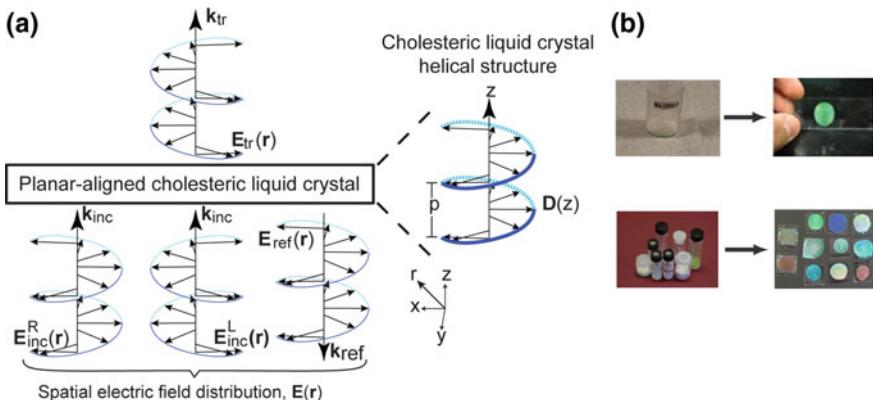


Fig. 4.30 (Reproduced with permission from [333]). **a** Circular polarized light propagation in a 1-D photonic bandgap CLC structure. **b** Preparation of photonic bandgap CLC structures from a powder oligomer (top) and monomeric LCs (bottom)

in a CLC mixture, and HTP is the helical twisting power of the chiral additive. For instance, with a chiral additive CB15 in the CB15/E7 mixture, $n_{av} \sim 1.6$ for this mixture, and $HTP \sim 7.3 \mu\text{m}^{-1}$ is the helical twisting power of the chiral additive in nematic liquid crystal E7. The actual stop band position relative to the fluorescence maximum of the emitter was further defined empirically by obtaining selective transmission curves of different samples using a spectrophotometer. After monomeric CLC preparation, a NQD solution of $\sim 1 \text{nM}$ concentration was mixed with monomeric CLC and the NQD solvent was evaporated.

For the oligomeric CLC powders [108, 355, 356], there is not such a simple relation between the concentration and λ_c . We found the right ratio R of components only empirically by mixing the different ratios of two oligomeric CLC powders with different λ_c by dissolving them in a solvent. By evaporating the solvent using a procedure of heating this solution in a vacuum inside a rotating retort, we obtained a new powder oligomer with an intermediate λ_c .

After selection of a proper concentration of components, monomeric CLC doped with emitters was placed between two cover glass slips with a standard planar alignment procedure. For planar alignment of oligomeric CLC, a cover glass slip with Wacker Chemie cyclosiloxane oligomeric powder [355] was placed on a hot plate and melted at $\sim 120^\circ\text{C}$ or higher. After a planar alignment with the second glass slip, the sample was slowly cooled into the glassy state preserving CLC order and planar alignment. For single-molecule fluorescence experiments some Wacker powders need to be purified. In many cases we purified CB15 and E7 as well using 0.3–5 μm particle filters.

We prepared resonance planar-aligned cholesteric structures for single emitters with fluorescence in visible and near-IR (Fig. 4.30b) using both cholesteric oligomeric powders (Fig. 4.30b, top) and monomeric nematic fluids with chiral additives (Fig. 4.30b, bottom). It is very important that because we used a high NA,

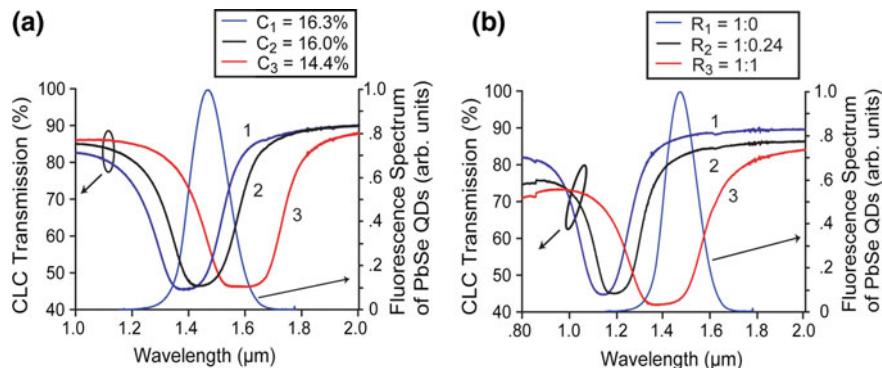


Fig. 4.31 (Reproduced with permission from [108]). CLC photonic bandgap selective transmission curves in unpolarized light for NQDs with fluorescence at optical communication wavelength ($1.5 \mu\text{m}$) for either monomeric (**a**) or oligomeric (**b**) CLC. Also shown are PbSe NQD fluorescence spectra at $1.5 \mu\text{m}$ that these transmission curves were turned to. **a** The weight concentrations of chiral additive CB15 in CB15/E7 mixture are $c_m = 16.3\%$ (curve 1), $c_m = 16.0\%$ (curve 2), and $c_m = 14.4\%$ (curve 3). **b** The weight ratio R of low- λ_c to high- λ_c Wacker oligomeric powders with $\lambda_c = 1.17 \mu\text{m}$ and $\lambda_c = 2.15 \mu\text{m}$ are: $R = 1:0$ (curve 1), $1:0.24$ (curve 2), and $1:1$ (curve 3)

oil immersion microscope objective ($\text{NA} = 1.35$) for single-emitter excitation, and a very short working distance ($\sim 200 \mu\text{m}$), we used cover glass slips with $\sim 120 \mu\text{m}$ thickness as the substrates for the liquid crystal cells. In this case of very fragile substrates, planar alignment using the buffering technique was challenging. We used a special mount in a buffering machine to fix these cover slips. For some experiments we used photoalignment. In some cases, a decent planar alignment on a small area was obtained by a simple unidirectional motion of the cover slips relative to each other. We used glass microspheres as spacers to fix the CLC cell thickness. Figure 4.31 shows selective transmission curves obtained in unpolarized light for SPS operating at optical communication wavelength $1.5 \mu\text{m}$ for fluid monomeric (a) and solid glassy oligomeric (b) CLCs. PbSe NQD fluorescence spectra at $1.5 \mu\text{m}$ are shown in both figures.

4.4.2 CLC Glassy Oligomeric Microcavities: NQD Fluorescence Enhancement and Antibunching

For SPS experiments we prepared 1-D photonic bandgap planar-aligned glassy CLC structures made of left-handed oligomeric CLC powder from Wacker Chemie doped with CdSeTe NQDs, (Qdot 800 ITK organic, Invitrogen, with fluorescence maximum at 790 nm [258, 333]). Doping was accomplished by heating the oligomer powder to $\sim 135^\circ\text{C}$ (the oligomer's melting temperature) and then mixing the melted CLC with NQDs dispersed in toluene, with heating allowed to continue until the toluene evaporated. Subsequently, cells were prepared using two

polyimide buffed glass coverslips with thickness $\sim 120 \mu\text{m}$. The CLC doped with NQDs was placed on a buffed coverslip and heated beyond the oligomer clearing temperature of 180°C . After the sample was cooled to $\sim 135^\circ\text{C}$, the second buffed coverslip was placed on the first and sheared along the direction of polyimide buffering. A slow cooling process back to a glassy (solid) state preserved the liquid crystal order. This resulted in a photonic bandgap microcavity with a center wavelength of 910 nm.

We excited the sample with cw, 633 nm laser light from a HeNe laser using a confocal fluorescence microscope with an oil-immersion, 1.35 NA objective. Figure 4.32a shows a confocal fluorescence microscope raster scan taken of a sample prepared using a $\sim 10 \text{ nM}$ concentration of NQDs dispersed in toluene, with the higher intensity spots indicating the location of fluorescing NQD in the glassy CLC microcavity. Focusing on a spot (circled in white on Fig. 4.32a, we checked for photon antibunching and obtained the coincidence histogram shown in Fig. 4.32b.

Figure 4.32b displays the histogram of coincidence counts $c(t)$ in blue, with $g^{(2)}(t)$ derived by normalizing $c(t)$. The measured $g^{(2)}(0)$ value from the fit shown in Fig. 4.32b (green, solid curve) is $g^{(2)}(0) = 0.382 \pm 0.037$. As $g^{(2)}(0) < 0.5$, this indicated that we have managed to isolate the fluorescence of a single quantum dot, serving as a source of antibunched light in a glassy CLC microcavity [258, 333].

By placing an achromatic quarter waveplate and linear polarizer in front of the spectrometer, we were able to filter for different handedness of circularly polarized (CP) fluorescence. The resulting fluorescence spectra of NQDs in the CLC microcavity can be seen in Fig. 4.33. NQD concentration was $\sim 1 \mu\text{M}$ in these experiments. Left-handed (LH) CP light experienced the photonic bandgap and therefore the black LHCP curve in Fig. 4.33 (curve 1) shows microcavity resonance, indicating that the LHCP light coupled to the cavity mode. The LHCP fluorescence had a center wavelength of 833 nm and a FWHM of 16 nm ($Q \sim 50$), as compared to a FWHM of 76 nm for RHCP fluorescence. The center wavelength of this resonance roughly matches the edge of the photonic stop band. The stop band is centered at 910 nm, as shown in Fig. 4.33 by the blue curve 3.

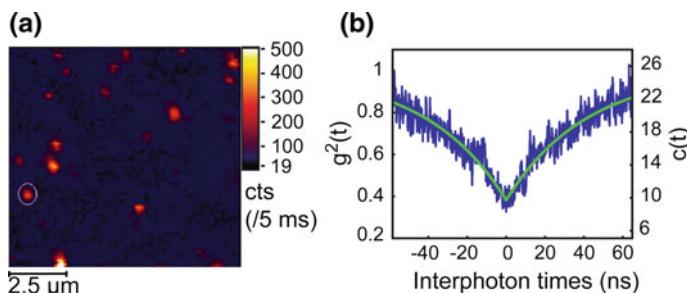


Fig. 4.32 (Reproduced with permission from [333]). **a** Confocal fluorescence microscopy image of single CdSeTe NQDs in a glassy CLC photonic bandgap microcavity. **b** Raw coincidence counts $c(t)$ (right-hand scale) and $g^{(2)}(t)$ (left-hand scale), showing antibunching (dip at $t = 0$)

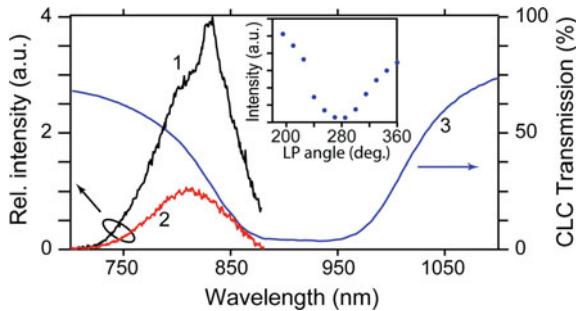


Fig. 4.33 (Reproduced with permission from [258, 333]). Circularly-polarized fluorescence resonance from NQDs doped in a glassy CLC microcavity. Curve 1: LHCP fluorescence spectrum of the NQDs with a resonance at 833 nm. Curve 2: RHCP fluorescence spectrum for the same NQDs. Curve 3: selective transmission of LHCP light through CLC microcavity. Inset: Dependence of resonance peak intensity on rotation of a linear polarizer (LP) after a fixed quarter wave plate. The quantum dot began bleaching while the measurement was made, so that the intensity at the second peak in the inset is less than at the first peak

The observed right-handed (RH) CP fluorescence is shown in red in Fig. 4.33 (curve 2) and was less intense due to not experiencing the CLC microcavity, showing no sign of line narrowing. The maximum intensity ratio between LHCP and RHCP was a factor of 4.9. To characterize the degree of circular polarization, the circular polarization dissymmetry factor g_e was used:

$$g_e = 2(I_L - I_R)/(I_L + I_R), \quad (4.5)$$

where I_L and I_R are intensities of LHCP and RHCP light, respectively. At the wavelength of the resonance shown in Fig. 4.33, $g_e = 1.3$. Note that for unpolarized light, $g_e = 0$, which we observed when NQDs were spin-coated on a bare glass slip.

4.4.3 CLC Monomeric Microcavities: Circular Polarized Fluorescence and Photon Antibunching

CLC Microcavities with NQDs

Using right handed, monomeric (fluid-like) CLC we obtained right-handed circular polarization of NQD fluorescence. Figure 4.34a shows a confocal microscope image of several single CdSe/ZnS NQDs fluorescing in a monomeric 1-D photonic bandgap CLC structure (E7 and CB15). A 532 nm laser with 6 ps pulse duration and 76 MHz pulse repetition rate was used in this experiment. Figure 4.34b (black and red solid curves) shows emission spectra for NQDs in a CLC microcavity for right-handed (black curve with high fluorescence intensity) and left-handed (red curve with low fluorescence intensity) circular polarizations [108, 333].

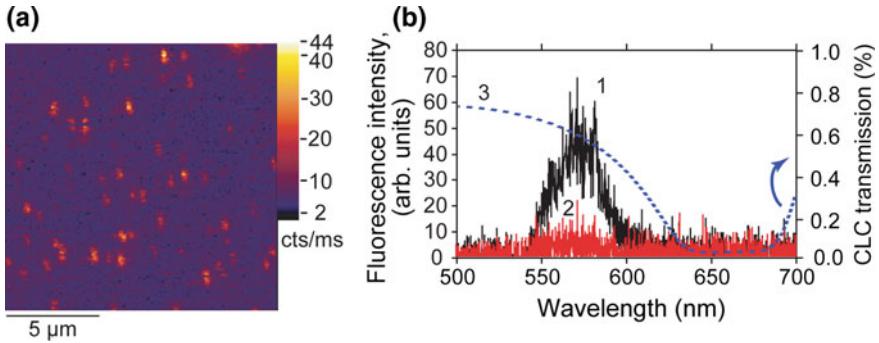


Fig. 4.34 (Reproduced with permission from [108, 333]). **a** Confocal fluorescence microscopy image of single CdSe NQDs in a monomeric CLC photonic bandgap microcavity. **b** Fluorescence spectrum of CdSe NQDs in the monomeric planar-aligned CLC host for two different circular polarizations of single photons (black curve of high intensity—right handed, red curve—left handed). CLC selective transmission curve is shown by a dashed blue line

The maximum intensity ratio between RHCP and LHCP was a factor of 5. The selective transmission curve of a CLC structure is shown as well (dashed blue curve). The degree of circular polarization measured by the dissymmetry factor was $g_e = -1.6$ (4.5) at 580 nm, showing right-handed circular polarization.

Figure 4.35a presents the $g^{(2)}(t)$ histogram at different interphoton times t . One sees that the peak at zero interphoton time is clearly smaller than any of the other peaks, which shows an antibunching property [$g^{(2)}(0) = 0.76 \pm 0.04$]. Figure 4.35b shows the Raman spectrum of undoped CLC (E7 and CB15). The peak near 580 nm is not a microcavity effect, because we observed the same features from unaligned CLC without a microcavity. This Raman spectrum of monomeric CLC explains the relatively high value of $g^{(2)}(0)$ in these experiments, but this antibunching histogram can be improved by using NQDs with a fluorescence spectrum outside the Raman spectrum of the CLC.

In order to avoid the monomeric CLC host Raman scattering, we selected CdSeTe NQDs with $\lambda_{\max} = 700$ nm and doped them into an CLC microcavity comprised of an E7 and CB15 mixture with a stop band edge at ~ 700 nm. When illuminating a single NQD, we obtained antibunching with $g^{(2)}(0) = 0.001 \pm 0.03$ (Fig. 4.35c). The spontaneous decay rate for these NQDs (20 MHz) is less than the laser excitation rate (76 MHz), so we cannot observe fluorescence excited by separate laser pulses as in Fig. 4.35a. Figure 4.35d shows the antibunching histogram for fluorescence of a CdSeTe NQD with $\lambda_{\max} = 700$ nm doped in the same E7 and CB15 CLC host illuminated with cw, 532 nm light (in difference with the pulsed 532 nm excitation used in Fig. 4.35c). We observed antibunched fluorescence with $g^{(2)}(0) = 0.11 \pm 0.06$, also showing that only single photons are emitted.

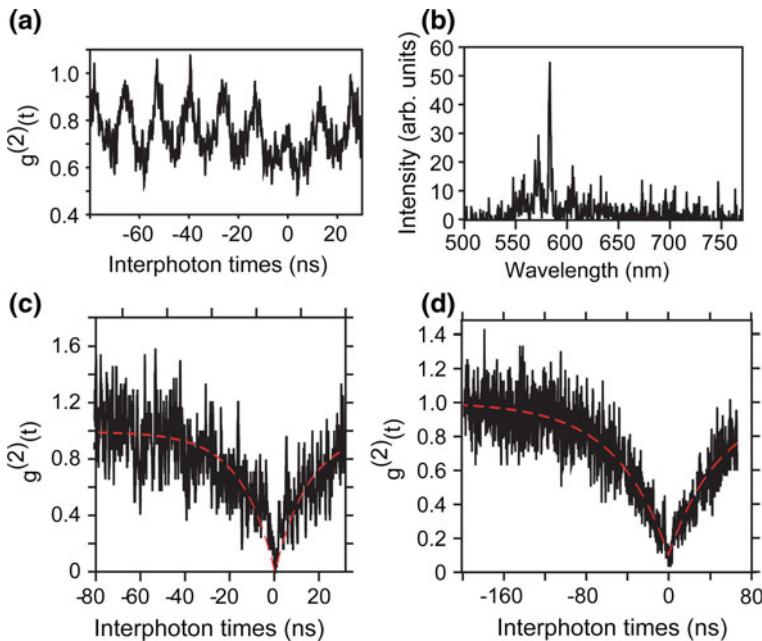


Fig. 4.35 (Reproduced with permission from [333]). **a** Histogram of coincidence counts of single CdSe NQD fluorescence ($\lambda_{\text{max}} = 580 \text{ nm}$) in a monomeric CLC host (E7 and CB15 mixture) under pulsed excitation. The dip at zero interphoton time indicates antibunching. **b** Raman spectra of E7 and CB15 mixture without NQDs showing maximum at a fluorescence wavelength $\sim 580 \text{ nm}$ of a CdSe NQD. **c, d** Antibunched fluorescence of CdSeTe NQDs ($\lambda_{\text{max}} = 700 \text{ nm}$) in a monomeric CLC host (E7 and CB15 mixture) using 532 nm, pulsed (see explanation in the text) and 532 nm, cw excitation

CLC Monomeric Microcavities with Nanocrystals Doped with Rare-Earth Ions

A monomeric CLC microcavity made from a mixture of E7 and CB15 was doped with rare-earth Er^{3+} and Yb^{3+} ions in 20–30 nm sized NaYF_4 nanocrystals with 20% Yb and 2% Er. When these ions were excited using a cw, 976 nm diode laser at incident powers of $\sim 500 \mu\text{W}$, we were able to observe upconverted fluorescence of Er^{3+} in the visible, as shown in Fig. 4.36a (red solid curve). The blue dashed curve shows the spectral transmission of CLC microcavity measured with unpolarized light.

Energy levels of Er^{3+} and Yb^{3+} and the upconversion excitation scheme under 980 nm excitation is shown in Fig. 4.36b. The emission lines observed were attributed to the transitions $^2\text{H}_{11/2}$, $^4\text{S}_{3/2} \rightarrow ^4\text{I}_{15/2}$ (green light) and $^4\text{F}_{9/2} \rightarrow ^4\text{I}_{15/2}$ (red light) of the Er^{3+} ions. The populations of upper levels in Er^{3+} occur due to an efficient energy transfer from the Yb^{3+} to the Er^{3+} . This fluorescence was measured to have a circular polarization dissymmetry factor of $g_e = -0.77$ at 680 nm [333]. The maximum intensity ratio between RHCP and LHCP was a factor of 2.25. These nanocrystals doped with rare-earth ions were prepared at the University at Buffalo

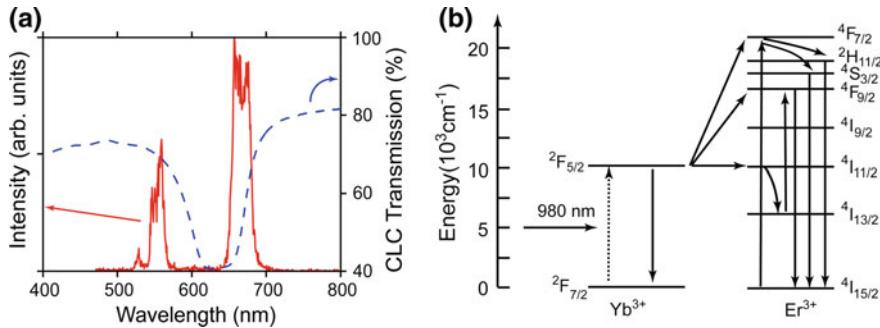


Fig. 4.36 (Reproduced with permission from [333]). **a** Red solid lines: Fluorescence spectrum of Er³⁺ ions doped in NaYF₄ nanocrystals dispersed in a chiral CLC microcavity (E7 and CB15). The blue dashed curve shows the spectral transmission of CLC microcavity measured with unpolarized light. **b** Energy levels of Er³⁺ and Yb³⁺ and upconversion excitation scheme under 980 nm excitation

(The Institute for Lasers, Photonics and Biophotonics). It should be noted, that Er³⁺ ions also have a fluorescence line at $\sim 1.5 \mu\text{m}$, which is important for fiber optical communication.

Single Photons from NV Center Nanodiamonds in a Cholesteric Liquid Crystal Host

We obtained photon antibunching from nitrogen vacancy color center in nanodiamond in a CLC (E7 and CB15) microcavity with a stop band centered at 725 nm, under cw, 514 nm argon ion laser excitation. The $g^{(2)}(0)$ value determined via fit was $g^{(2)}(0) = 0.74 \pm 0.08$ [333]. These nanodiamonds were purchased from Microdiamant AG, Switzerland. They are agglomerate free, and $\sim 30\text{--}50 \text{ nm}$ in diameter.

4.4.4 Single-Emitter Alignment, Single-Photon Sources with Definite Linear Polarization, and Some Plasmonics Applications of Liquid Crystals

Another intriguing application of liquid crystals as hosts for single emitters is using the planar-alignment of liquid crystals (nematic or smectic) to provide doped molecular dipoles of emitters with definite alignment along a preferred direction for efficient excitation [122, 337, 339, 341, 342]. This allows for the creation of a single-photon source with fluorescence of definite linear polarization. For these experiments, polymeric, oligomeric, and monomeric LCs were used [122, 337, 339, 341, 342]. It should be also mentioned that single terrylene dye molecules doped into linear low-density polyethylene have been oriented by tensile deformation of

the matrix without using any LC [357]. In measurements on ensembles at ambient and on single terrylene molecules at cryogenic temperature, strong orientation along the stretching direction was observed by polarization-resolved confocal fluorescence microscopy.

Linear Polarization of Single Photons from a Dye-Doped Glassy Nematic LC Oligomer

A glassy nematic liquid crystal oligomer synthesized by S. H. Chen's group (University of Rochester) was doped with DiIIC₁₈(3) (DiI) dye molecules from Molecular Probes [339]. Planar-aligned layers of this doped liquid crystal host of ~100 nm thickness were prepared using photoalignment of the liquid crystal molecules. Photoalignment was performed by spin-coating a Staralign-2100 linearly photopolymerizable polymer (Rolic Technologies Ltd.) to a cleaned cover glass slip, which was then cured at ~135 °C. This film was then irradiated by polarized UV light for 10–15 min, with further irradiation used to bleach the polymer impurity fluorescence [339].

An oligomer solution doped with dye and diluted in chloroform was subsequently spin-coated onto these Staralign coated glass slips. After the chloroform evaporated, we heated the sample to ~80 °C, slightly above where the oligomer transition to a nematic state occurs, after which the sample was slowly cooled to a glassy state, preserving the planar-aligned nematic order.

To characterize prepared samples, we used a confocal microscope (WITec alpha-SNOM in confocal transmission mode), exciting the sample using a cw, 532 nm Nd:YAG laser excitation. Single-photon counting avalanche photodiodes were used as photodetectors for confocal fluorescence scans, with a polarizing beamsplitter used so that each photodetector collected light of orthogonal polarizations.

DiI dye molecules doped in a planar-aligned liquid crystal host tend to fluoresce with polarization perpendicular to the alignment of the liquid crystal. It can be explained by the molecular structure of these molecules. As illustrated in Fig. 4.37a, it is likely that two alkyl chains of these molecules orient themselves parallel to the rod-like molecules of the nematic liquid crystal host. The absorbing and emitting dipoles, however, are parallel to the bridge between these alkyl chains. Therefore, these dipoles end up oriented perpendicular to the direction of liquid crystal alignment, hence the fluorescence having a polarization orthogonal to the alignment of the liquid crystal. The micrographs in Fig. 4.37b, c show confocal microscope raster scan images of DiI single-molecule fluorescence for polarization components perpendicular (b) and parallel (c) to the nematic alignment direction.

38 single molecules were identified by the peak pixel intensity values (with background subtraction) and the degree of polarization, ρ , was determined for each molecule by comparison of the “perpendicular” and “parallel” polarization images. We define the degree of polarization [121, 122, 341, 342]

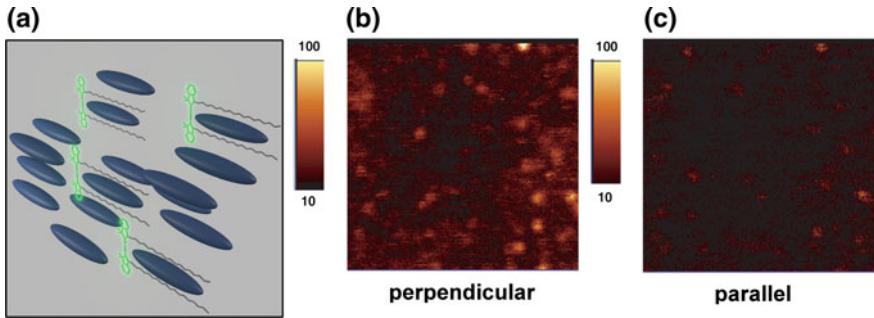


Fig. 4.37 (Reproduced with permission from [333]). **a** Schematic view of a DiI dye molecules in a nematic liquid crystal host. The long axis (alkyl chains) of the DiI molecules tend to orient themselves along the rod-like nematic molecules, while the dipole, which is parallel to the bridge between the alkyl chains, orients perpendicular to the direction of liquid crystal alignment. **b, c** Confocal fluorescence microscopy images of DiI dye single-molecule fluorescence in a planar-aligned, glassy, nematic liquid crystal host ($10 \mu\text{m} \times 10 \mu\text{m}$ raster scan). Scale shows counts/pixel

$$\rho = (I_{\text{par}} - I_{\text{perp}}) / (I_{\text{par}} + I_{\text{perp}}), \quad (4.6)$$

where I_{par} and I_{perp} are, respectively, the fluorescence intensities parallel and perpendicular to the direction of alignment. Note that the only difference of ρ with the degree of polarization $|\rho|$, usually used in the literature, e.g., in [121, 122, 341, 342] is that our definition allows inclusion of the direction of linear polarization as part of the measure, depending on whether ρ is positive or negative.

The different ρ values that we found are histogrammed in Fig. 4.38a. A clear asymmetry in Fig. 4.38a demonstrates a preference towards fluorescence polarized *perpendicular to sample alignment*. This is in contrast to the expected fluorescence from an unoriented sample, which would yield a symmetric ρ histogram relative to $\rho = 0$.

To confirm these results, additional experiments were made with spectrophotometer for polarization perpendicular and parallel to the sample alignment, this time having used a sample of planar-aligned glassy nematic liquid crystal oligomer doped with the same DiI molecules, but of high concentration (more than 1% by weight). The value of ρ was calculated by comparing the peak intensities of the curves (see Fig. 4.38b), yielding $\rho = -0.5$, that also indicates a clear preference for fluorescence with definite linear polarization *perpendicular to the sample alignment*.

Using the same DiI dye doped in monomeric planar-aligned nematic liquid crystal E7, we carried out measurements on the dependence of dye fluorescence intensity as the linear polarization of the excitation 532 nm laser beam was rotated over 360° (Fig. 4.38c). This figure shows a clear dependence of the fluorescence intensity of an ensemble of DiI molecules on the exciting angle of polarization, where the maximum fluorescence occurred when the DiI molecules were excited by

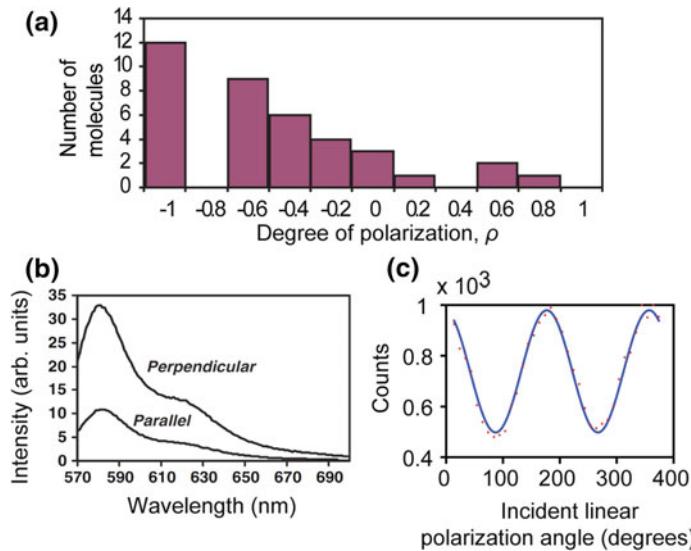


Fig. 4.38 (Reproduced with permission from [333]). **a** Asymmetric histogram of polarization measure ρ from 38 different DiI dye molecules in a planar-aligned glassy nematic LC host showing linear polarization distribution of single dye molecules. Negative values in our definition of ρ correspond to polarization perpendicular to alignment direction. For random polarization a histogram will be symmetrical. **b** Polarized fluorescence of DiI dye in planar aligned glassy nematic LC at *high concentration* of dye. Curves show fluorescence spectrum with polarization perpendicular and parallel to the host alignment direction. **c** (From [79]). Intensity of DiI fluorescence in planar-aligned E7 nematic LC as linear polarization of excitation laser light was rotated over 360° . The zero angle of polarization corresponds to the direction perpendicular to liquid crystal alignment

light with linear polarization perpendicular to the alignment of the liquid crystal [333]. In this experiment the LC alignment was accomplished by simple shearing the two cover glass slips in one direction with a drop of dye-doped LC inside.

Alignment of Dot-in-Rods (DR) with Liquid Crystals

Using LC photoalignment, colloidal semiconductor CdSe/CdS core-shell DRs dispersed in LC polymer were aligned in a preferred direction [122]—see Fig. 4.39. In [122], DRs were referred to as nanorods (NRs). The aligned NRs with an aspect ratio of 5:1 (4 nm width and 20 nm length) exhibited linearly polarized emission near 586 nm with a degree of polarization of 0.62. The order parameter S of aligned NRs area was 0.87, where $S = \langle 2 \cos^2 \theta - 1 \rangle$, and θ is the relative angle of the NRs to the main axis. $S = 1$ defines perfectly aligned NRs, and $S = 0$ defines a completely random system.

The maximum degree of polarization for a *single* DR was ~ 0.8 [121]. That work used CdSe/CdS DRs with aspect ratio ~ 19 . It has been reported that higher aspect

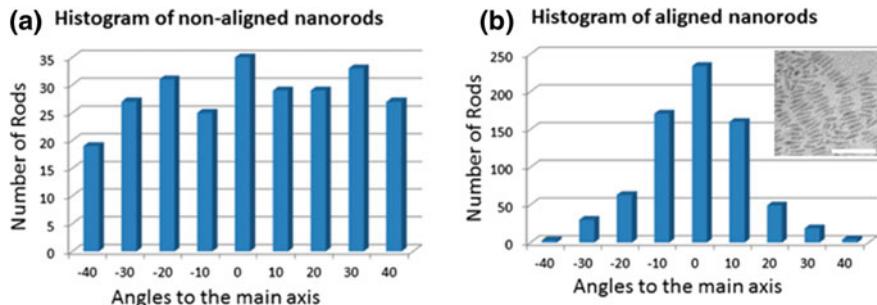


Fig. 4.39 (Reproduced with permission from [122]). **a** The histogram of nonaligned NRs, showing the directional distribution of NRs, with the 0° bar indicating the number of rods within an angle between -5° and 5° to the main axis. The order parameter for this area is 0.22. **b** The histogram of the aligned NRs. The order parameter for this area is 0.87 according to the directional distribution of NRs. Insert in (b) shows TEM image of NRs aligned within the LC polymer matrix (white scale bar is 50 nm)

ratios may adversely affect the emitter photoluminescence quantum yield. The NRs in [122] had a photoluminescence quantum yield of 84%.

In [341, 342], the alignment of CdSe/CdS DRs in a preferred direction was made using linear defects (oily streaks) in smectic LC thin films (8CB) [358]. After standard rubbing procedure on the polymer coated substrate with defined rubbing direction, linear defects in smectic film (see Fig. 4.40a, b) appeared in the direction perpendicular to the rubbing direction. When the DRs are localized in oriented oily streaks, their fluorescence polarization is maximum in a well-defined direction parallel to the stripes (perpendicular to rubbing direction). It was shown that the DRs are confined parallel to these linear defects. This technique also prevents aggregation of DRs. DRs were constituted of a CdSe core with a diameter of 2.9 nm and were surrounded by a rod-like CdS shell with average length 23 nm and had a total thickness \sim 7 nm. Oily streak defects visualized by optical microscopy appeared in 8CB films of thickness varying between 100 and 300 nm. The histogram of Fig. 4.40c shows linear polarized emission of aligned DRs.

Liquid Crystals and Some Plasmonics Applications

Liquid crystals' capability of changing refractive index with temperature or electric field can find applications in tunable SPSs. In [359] the results are presented on calculations of plasmonic resonance shifts of gold and silver bowtie nanoantennas in liquid crystals, and experimental fluorescence enhancement of NV center nanodiamonds in liquid crystals with silver nanorods. Additionally, liquid crystals can increase the life of silver plasmonic devices protecting them from ambient air.

In [360], a composite material containing well-known commercially available liquid crystals was developed with reconfigurable fluorescence properties, in which both the intensity and the decay rates of fluorescence have been controlled by electric fields (several volts). This composite material combines rod-like gold-

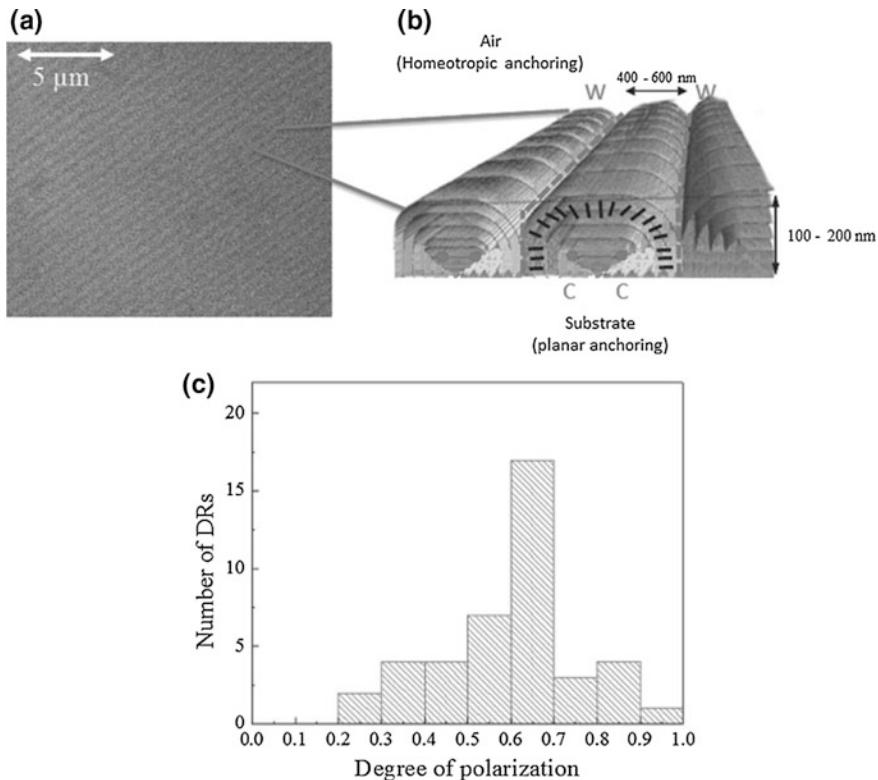


Fig. 4.40 (Reprinted with permission from [341]). **a** Oily streaks observed by optical microscopy on a sample 120 nm thick, viewed from the top. **b** Scheme of the oily streaks with the smectic layers schematically shown in 3D and the molecules' orientation indicated in black. The centers of curvature are indicated by dots (marked as C for the central hemicylinder), and the walls between the hemicylinders, W, are highlighted with dashed lines. **c** Histogram of the measured degrees of polarization for 42 DRs in 8CB liquid crystal films of thickness 100 nm

silica-dye (GSD) colloidal nanostructures (\sim 100–500 nm length) and a liquid crystal host medium that imposes orientational ordering on these nanorods by planar alignment of liquid crystal cells. Dye molecules are randomly distributed within silica shells of these core-shell GSD nanorods.

4.5 Fiber Integrated SPSs

Tapered optical fibers were used for optical coupling of NV nanodiamonds with resonator modes of a toroid [273] and with single-mode tapered diamond waveguides with NV centers [361]. Color-center nanodiamonds have been placed on the

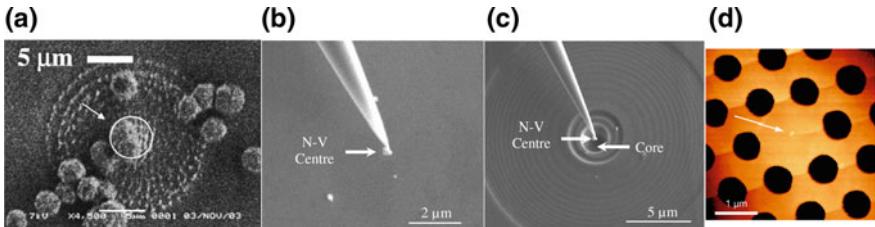


Fig. 4.41 **a** (Reproduced with permission from [154]). SEM image of nanodiamonds grown on an optical fiber facet. **b** (Reproduced with permission from [362]). SEM image of a single nanodiamond (~ 300 nm) being removed from the marked silicon substrate. **c** SEM image of the tapered probe with the nanodiamond attached and positioned above the optical fiber facet. **d** (Reproduced with permission from [156]): AFM image of the 30-nm-nanodiamond placed in the middle of the core (marked by an arrow) of a photonic crystal fiber via an AFM pick-and-place technique

facets of optical fibers [154–156, 362]; the first commercial SPS (Quantum Communication Victoria) was based on this technology. For instance, in [154], NV center nanodiamonds were grown on a fiber facet using chemical vapor deposition (Fig. 4.41a). The nanomanipulation of single preselected nanodiamonds was achieved using a custom tapered optical fiber probe covered by a ~ 10 nm carbon layer in a standard SEM with a nanomanipulator [362], see Fig. 4.41b, c. When approaching a probe tip to the preselected nanodiamond on a marked Si substrate (Fig. 4.41c) with nanodiamonds within a SEM, the nanodiamond appears to energetically favor the conductive probe tip and in some cases “jumps” from the substrate onto the tip with a high yield of $\sim 90\%$ [362]. The selected nanodiamond was then transferred on to an optical fiber facet. The fiber facet was coated with a 20 nm-thick conductive carbon layer to allow the nanodiamond to be manipulated onto the surface once contact was made between the nanodiamond and the substrate. The process was successful $\sim 70\%$ of the time (Fig. 4.41c). Because of fluorescence background from fiber-core impurities, photon antibunching was obtained for nanodiamonds placed outside the core [362].

In [156] the nanodiamond was deposited via pick-and-place manipulation [155] by an AFM tip in the core region of a photonic crystal fiber with a diameter of 90 μm , a core region of 1.5 μm , and a length of about 10 cm. Figure 4.41d shows an AFM micrograph of this fiber core with the nanodiamond on the facet placed close to the center of the core. The nanodiamond was picked up by an AFM tip repeatedly pressing the tip down on the nanodiamond with a force on the order of magnitude of 0.1 μN [156]. After the AFM tip was brought manually in a position atop the fiber cladding, the nanodiamond was placed on the cladding and transferred to the core using the same technique. Photon antibunching with $g^{(2)}(0) < 0.5$ was reported for both cw and pulsed laser excitation.

4.6 Conclusion and Outlook

Advances in room-temperature SPSs are closely related with the development of nanophotonic technology, starting from selection of single emitters to create antibunched photons, to engineering the required properties of emitted photons by photonic/plasmonic/metamaterial nanostructures. In today's implementations we can find SPSs with one or two aspects of the ideal implementation (see Sect. 4.1.3), but performance of other parameters is often compromised [69]. In particular, the SPS efficiency (as defined in Sect. 4.3.1) is usually not more than several percent. The selection of a given system depends on a particular practical application of the SPS. For example, quantum computation protocols with antibunched photons require all photons to be indistinguishable from one to another (a Fourier-transform limited emission); this allows the possibility of quantum interference, which for single-emitter systems has been only achieved under cryogenic temperatures [363–375]. For secure quantum communication, only *room-temperature* SPSs are practical. SPSs operated at room-temperature have the advantages of compactness and less power consumption (no cryogenic systems) that are critical for quantum communication with satellites. For modern QKD systems, SPSs should be bright, efficient, on-demand, and have definite polarization, high repetition rate (at least GHz) and low jitter (100 ps) at required wavelengths (750–1600 nm for a free-space QKD satellite-to-ground configuration, and 1.3 and 1.55 μm for fiber-optical communication). For daylight free-space quantum communication and in fiber optical communication, narrow spectral band transmission filters can be used with bright room-temperature SPSs, e.g., atomic filters with GHz transmission bandwidth [376, 377] or microcavities providing a linewidth of a few GHz [259–262, 378].

It should be noted that with development of QKD technology from simple to more complex systems, SPSs will be necessary elements of secure quantum communication networks providing higher security and key rates than the more convenient and cheap decoy state protocol [379]. In device-independent QKD [379–381] and long-distance (several hundreds or thousands of kilometers) quantum communication based on quantum repeaters [11–14, 17], SPSs could play a key role [379]. Both these applications require entanglement, which can be created using two on-demand sources of indistinguishable antibunched photons and a beamsplitter [379]. In [381] it was shown that using SPSs in the suggested entanglement swapping procedure for device-independent QKD increases the achievable key rate by several orders of magnitude, in comparison with the usual sources of entangled indistinguishable photons based only on spontaneous parametric downconversion.

The principle of quantum repeaters is based on repeated entanglement swapping between neighboring links. The entanglement has to be stored in each link and processed many times to be swapped between adjacent links until it is extended over the desired distance [11, 12]. Quantum repeater protocols based on SPSs [10, 382] eliminate the fundamental errors due to double-photon emission which limited

the performance of previous protocols using either nonclassical correlations between atomic excitations and emitted photons [13] or photon pair sources (including spontaneous parametric down conversion) [383]. The entanglement distribution rate with SPSs in [382] is three orders of magnitude greater than that in [13].

For quantum repeater protocols, single photons should be indistinguishable, with SPS efficiency higher than 70% and with a negligible double-photon emission (less than 10^{-4} [10, 379]). They also need to be compatible with the quantum memory wavelength and bandwidth (hundreds of MHz). Photon indistinguishability at room temperature is a very challenging task, although in [384] a high degree of indistinguishability was predicted using the example of two typical room-temperature single emitters within photonic crystal and fiber-based microcavities. A special regime was found in which the broad spectrum of the single emitter was funneled into a narrow cavity resonance. In [385], an ultra-small mode volume plasmonic-Fabry-Perot hybrid cavity with a SiV nanodiamond is discussed towards a room temperature indistinguishable SPS.

Finally, this review outlines the state-of-the-art in development of *room-temperature* SPSs. As a summary of Sect. 4.2, Table 4.2 provides comparisons between the main stable single emitters for room-temperature sources of single photons. See Sect. 4.2.4 for other single emitters for room-temperature SPSs.

Table 4.2 Comparison of photostable single emitters for SPS applications

Single emitter	Advantages	Disadvantages
Color-center diamond	<ul style="list-style-type: none"> • NV nanodiamonds are on the market; • Stability against bleaching; • SiV, Cr and NE8 centers have narrow lines 	<ul style="list-style-type: none"> • NV centers have wide fluorescence spectrum; • Difficult to find a single color center with $g^{(2)}(0) \sim 0$ • SiV center has lower stability and brightness
Giant nanocrystal quantum dots and dot-in-rods	<ul style="list-style-type: none"> • Relatively easy to prepare; • Relatively stable; • Fluorescence at both visible and optical communication wavelengths; • Possibility of alignment of dot-in-rods to create definite polarization of single photons 	<ul style="list-style-type: none"> • Less stable than color-center diamonds and rare-earth ions; • Wide spectral lines
Rare-earth ions in nanocrystals	<ul style="list-style-type: none"> • Emitters with highest stability (no bleaching); • Variety of ions with fluorescence wavelengths both in visible and at optical communication wavelengths; 	<ul style="list-style-type: none"> • Forbidden transitions (including 1.55 μm) have low oscillator strengths and long (millisecond) fluorescence lifetimes; • Allowed transitions have wide spectra and require two-photon excitation of specific wavelengths

In this review we also focused on the methods of enhancing spontaneous emission and efficient photon extraction using nano/microcavities, plasmonic nanoantennas and metamaterials. Each type of nanostructure for tailoring single-emitter fluorescence properties has advantages as well as challenges. For instance, a highest Purcell factor with increasing emitter radiative decay rate ($\sim 1,000$ times) [285, 305] and 2,300-fold enhancement in the total fluorescence intensity [286] were obtained with metal plasmonic nanoantennas. This was done by balancing the fluorescence enhancement and quenching by selecting the proper nanoantenna gap length. Such achievements make practical to use for SPSs even forbidden transitions (including $1.55\text{ }\mu\text{m}$ wavelength) of the rare-earth ions doped in nanocrystals.

Metamaterials, even with nanopatterning, do not yet provide such record decay rate increases and photon extraction efficiencies as do plasmonic nanoantennas. But for applications that require broad spectral band SPSs (such as when using quantum dots with variable emission wavelengths), they can be a choice. Metamaterials for fiber-optical communication wavelengths have been developed. It is also important that for metamaterials, the placement of a single emitter does not require high precision.

For narrow spectral bandwidth, micro/nanocavities should be used with single emitters. It is also important to place a single emitter in the position of the maximum nanocavity field. This can be challenging when using nanocrystals of several or tens of nanometers in size. As was mentioned earlier, single photons with a high degree of indistinguishability can be created within microcavities, even at room temperature [384, 385], although this goal is very challenging. Tunable liquid crystal cholesteric photonic bandgap microcavities and alignment of emitters with liquid crystals can also find applications in SPSs with definite circular or linear polarizations used in a BB84 QKD protocol.

To develop SPSs with desired parameters for QKD and operating at room temperature, hybrid photonic/plasmonic/metamaterial structures may be the best choice. Indistinguishable photons require optical micro/nanocavities, but plasmonic structures provide the highest Purcell factor and enhancement of total fluorescence intensity. In [385, 386] such hybrid cavity/plasmonic systems are modelled. In [387] a hybrid cavity SPS on a silicon chip was experimentally realized with a NV center in nanodiamond with up to 42-fold decay rate enhancement of the spontaneous emission at the cavity resonance. For fiber-optical communications, it is desirable to integrate SPSs into microcavities milled in fiber facets, as was made in [259–261] for microcavities (Fig. 4.19). Plasmonic nanoantenna on a fiber facet was used in [388] and a metamaterial on a fiber facet was prepared in [389]. Compact room-temperature SPSs are also necessary for on-chip-operation [71, 299]. Plasmonic materials for telecom wavelengths are discussed in [390].

One important direction that can facilitate the telecom wavelength problem for room-temperature SPSs with “ideal” properties for QKD is spectral translation to and from visible and telecom wavelengths, via nonlinear optical processes. This technique [391–393] gives the advantage of using Si detectors, and preserves photon antibunching. Reference [394] modeled in-fiber wavelength conversion in

the spectral range 400–1700 nm, using intermodal four-wave mixing (Bragg scattering) between orbital angular momentum (OAM) fiber modes. Using OAM beams is another interesting direction in QKD. This method expands current capabilities of QKD protocols by transmitting several bits per single photon in the high-dimensional state space associated with the transverse mode structure. Conventional QKD systems relying on polarization encoding not only limit the amount of information that can be sent per photon, but also place tight bounds on the error rates that such a system can tolerate. QKD systems based on spatial-mode encoding using OAM better resist eavesdropping attack in addition to having an increased information capacity [395–398].

After this chapter submission new important papers appeared. For instance, on the calculation of two-cavity coupled system for efficient indistinguishable SPS at room temperature [399], on bright (80% quantum efficiency) tin (Sn) vacancy color center in diamond [400], and up to 35 million photon counts per seconds from 20 nm NV center nanodiamonds coupled to silver nanopatch antennas [401] with observation of photon antibunching within nanoantennas. The material platform for on-chip integration of quantum photonic systems is reviewed in [402].

Acknowledgements S. G. L. acknowledges support from the US National Science Foundation grants EEC 1343673 and ECC 0420888. L. J. B. was supported by the Office of the Secretary of Defense ARAP QSEP program; and AFOSR, “Single Color Center Engineering”. We also thank the following publishers for permission to reproduce their materials: Nature Publishing Group, Taylor & Francis Group, American Chemical Society, American Physical Society, Optical Society OSA, Elsevier, John Wiley and Sons, Institute of Physics, American Institute of Physics, and Turpion.

References

1. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, Quantum cryptography. *Rev. Mod. Phys.* **74**, 145–195 (2002)
2. V. Scarani, H. Bechmann-Pasquinucci, N.J. Cerf, M. Dušek, N. Lütkenhaus, M. Peev, The security of practical quantum key distribution. *Rev. Mod. Phys.* **81**, 1301–1350 (2009)
3. L. Mandel, E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995)
4. D.F. Walls, Evidence for the quantum nature of light. *Nature* **280**, 451–454 (1979)
5. H. Paul, Photon antibunching. *Rev. Mod. Phys.* **54**, 1061–1102 (1982)
6. C.H. Bennett, G. Brassard, Quantum cryptography: public key distribution and coin tossing, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, Bangalore, India, 1984), p. 175
7. N. Lütkenhaus, Security against individual attacks for realistic quantum key distribution. *Phys. Rev. A* **61**, 52304 (2000)
8. W.-Y. Hwang, Quantum key distribution with high loss: toward global secure communication. *Phys. Rev. Lett.* **91**, 57901 (2003)
9. H.-K. Lo, X. Ma, K. Chen, Decoy state quantum key distribution. *Phys. Rev. Lett.* **94**, 230504 (2005)
10. N. Sangouard, C. Simon, J. Minář, H. Zbinden, H. de Riedmatten, N. Gisin, Long-distance entanglement distribution with single-photon sources. *Phys. Rev. A* **76**, 50301 (2007)

11. H.-J. Briegel, W. Dür, J.I. Cirac, P. Zoller, Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998)
12. H.-J. Briegel, S.J. van Enk, J.I. Cirac, P. Zoller, D. Bouwmeester, J.-W. Pan, M. Daniell, H. Weinfurter, A. Zeilinger, V. Vedral, M.B. Plenio, P.L. Knight, Quantum networks and multi-particle entanglement, in *The Physics of Quantum Information*, ed. by D.D. Bouwmeester, P.A. Ekert, P.A. Zeilinger (Springer, Berlin Heidelberg, 2000), pp. 191–220
13. L.-M. Duan, M.D. Lukin, J.I. Cirac, P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics. *Nature* **414**, 413–418 (2001)
14. N. Sangouard, C. Simon, H. de Riedmatten, N. Gisin, Quantum repeaters based on atomic ensembles and linear optics. *Rev. Mod. Phys.* **83**, 33–80 (2011)
15. A.I. Lvovsky, B.C. Sanders, W. Tittel, Optical quantum memory. *Nat. Photonics* **3**, 706–714 (2009)
16. C. Simon, M. Afzelius, J. Appel, A.B. de la Giroday, S.J. Dewhurst, N. Gisin, C.Y. Hu, F. Jelezko, S. Kroll, J.H. Muller, J. Nunn, E. Polzik, J. Rarity, H. de Riedmatten, W. Rosenfeld, A.J. Shields, N. Skold, R.M. Stevenson, R. Thew, I. Walmsley, M. Weber, H. Weinfurter, J. Wrachtrup, R.J. Young, Quantum memories. A review based on the European Integrated Project “Qubit Applications (QAP)”. *Eur. Phys. J. D.* **58**, 1–22 (2010)
17. J.I. Cirac, P. Zoller, H.J. Kimble, H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78**, 3221–3224 (1997)
18. X. Maître, E. Hagley, G. Nogues, C. Wunderlich, P. Goy, M. Brune, J.M. Raimond, S. Haroche, Quantum memory with a single photon in a cavity. *Phys. Rev. Lett.* **79**, 769–772 (1997)
19. E. Knill, R. Laflamme, G.J. Milburn, A scheme for efficient quantum computation with linear optics. *Nature* **409**, 46–52 (2001)
20. T.B. Pittman, B.C. Jacobs, J.D. Franson, Demonstration of nondeterministic quantum logic operations using linear optical elements. *Phys. Rev. Lett.* **88**, 257902 (2002)
21. T.B. Pittman, B.C. Jacobs, J.D. Franson, Probabilistic quantum logic operations using polarizing beam splitters. *Phys. Rev. A* **64**, 62311 (2001)
22. A. Shields, Quantum logic with light, glass, and mirrors. *Science* **297**, 1821–1822 (2002)
23. G. Brida, M. Genovese, M. Gramegna, Twin-photon techniques for photo-detector calibration. *Laser Phys. Lett.* **3**, 115–123 (2006)
24. S. Buckley, K. Rivoire, J. Vučković, Engineered quantum dot single-photon sources. *Rep. Prog. Phys.* **75**, 126503 (2012)
25. M. Felle, J. Huwer, R.M. Stevenson, J. Skiba-Szymanska, M.B. Ward, I. Farrer, R.V. Penty, D.A. Ritchie, A.J. Shields, Interference with a quantum dot single-photon source and a laser at telecom wavelength. *Appl. Phys. Lett.* **107**, 131106 (2015)
26. D. Fattal, K. Inoue, J. Vučkovic, C. Santori, G.S. Solomon, Y. Yamamoto, Entanglement formation and violation of Bell’s Inequality with a semiconductor single photon source. *Phys. Rev. Lett.* **92**, 37903 (2004)
27. M.J. Stevens, Photon statistics, measurements, and measurements tools, in *Single-Photon Generation and Detection: Physics and Applications*, ed. by A. Migdall, S.V. Polyakov, J. Fan, J.C. Bienfang (Academic Press, 2013), pp. 25–68
28. R.J. Glauber, The quantum theory of optical coherence. *Phys. Rev.* **130**, 2529–2539 (1963)
29. R.H. Brown, R.Q. Twiss, Correlation between photons in two coherent beams of light. *Nature* **177**, 27–29 (1956)
30. R.J. Glauber, Coherent and incoherent states of the radiation field. *Phys. Rev.* **131**, 2766–2788 (1963)
31. M. Ehrenberg, R. Rigler, Rotational Brownian motion and fluorescence intensify fluctuations. *Chem. Phys.* **4**, 390–401 (1974)
32. H.J. Carmichael, D.F. Walls, A quantum-mechanical master equation treatment of the dynamical Stark effect. *J. Phys. B: At. Mol. Phys.* **9**, 1199 (1976)
33. H.J. Kimble, L. Mandel, Theory of resonance fluorescence. *Phys. Rev. A* **13**, 2123–2144 (1976)

34. C. Cohen-Tannoudji, Atoms in strong resonance fields, in *Frontiers in Laser Spectroscopy, Les Houches Summer School Proceedings*, Session XXVII, July 1975, ed. by R. Balian, S. Haroche, S. Liberman (North-Holland, Amsterdam, 1977), pp. 1–104
35. H.J. Kimble, M. Dagenais, L. Mandel, Photon antibunching in resonance fluorescence. *Phys. Rev. Lett.* **39**, 691 (1977)
36. H.J. Kimble, M. Dagenais, L. Mandel, Multiatom and transit-time effects on photon-correlation measurements in resonance fluorescence. *Phys. Rev. A* **18**, 201–207 (1978)
37. M. Dagenais, L. Mandel, Investigation of two-time correlations in photon emissions from a single atom. *Phys. Rev. A* **18**, 2217–2228 (1978)
38. J.D. Cresser, J. Häger, G. Leuchs, M. Rateike, H. Walther, Resonance fluorescence of atoms in strong monochromatic laser fields, in *Dissipative Systems in Quantum Optics*, ed. by P.R. Bonifacio (Springer, Berlin, Heidelberg, 1982), pp. 21–59
39. P. Kask, P. Piksarv, Ü. Mets, Fluorescence correlation spectroscopy in the nanosecond time range: photon antibunching in dye fluorescence. *Eur. Biophys. J.* **12**, 163–166 (1985)
40. Ü. Mets, Antibunching and rotational diffusion in FCS, in *Fluorescence Correlation Spectroscopy, Theory and Applications*, vol. 65, Chemical Physics, ed. by R. Rigler, E.S. Elson (Springer, Berlin, Heidelberg, 2001), pp. 346–359
41. J.C. Bergquist, R.G. Hulet, W.M. Itano, D.J. Wineland, Observation of quantum jumps in a single atom. *Phys. Rev. Lett.* **57**, 1699–1702 (1986)
42. F. Diedrich, H. Walther, Nonclassical radiation of a single stored ion. *Phys. Rev. Lett.* **58**, 203–206 (1987)
43. W.E. Moerner, L. Kador, Optical detection and spectroscopy of single molecules in a solid. *Phys. Rev. Lett.* **62**, 2535–2538 (1989)
44. M. Orrit, J. Bernard, Single pentacene molecules detected by fluorescence excitation in a p-terphenyl crystal. *Phys. Rev. Lett.* **65**, 2716–2719 (1990)
45. E. Betzig, R.J. Chichester, Single molecules observed by near-field scanning optical microscopy. *Science* **262**, 1422–1425 (1993)
46. T. Basché, W.E. Moerner, M. Orrit, H. Talon, Photon antibunching in the fluorescence of a single dye molecule trapped in a solid. *Phys. Rev. Lett.* **69**, 1516–1519 (1992)
47. F. De Martini, G. Di Giuseppe, M. Marrocco, Single-Mode Generation of quantum photon states by excited single molecules in a microcavity trap. *Phys. Rev. Lett.* **76**, 900–903 (1996)
48. W. Patrick Ambrose, P.M. Goodwin, J. Enderlein, D.J. Semin, J.C. Martin, R.A. Keller, Fluorescence photon antibunching from single molecules on a surface. *Chem. Phys. Lett.* **269**, 365–370 (1997)
49. S.C. Kitson, P. Jonsson, J.G. Rarity, P.R. Tapster, Intensity fluctuation spectroscopy of small numbers of dye molecules in a microcavity. *Phys. Rev. A* **58**, 620 (1998)
50. C. Brunel, B. Lounis, P. Tamarat, M. Orrit, Triggered source of single photons based on controlled single molecule fluorescence. *Phys. Rev. Lett.* **83**, 2722 (1999)
51. B. Lounis, W.E. Moerner, Single photons on demand from a single molecule at room temperature. *Nature* **407**, 491–493 (2000)
52. L. Fleury, J.M. Segura, G. Zumofen, B. Hecht, U.P. Wild, Nonclassical photon statistics in single-molecule fluorescence at room temperature. *Phys. Rev. Lett.* **84**, 1148–1151 (2000)
53. F. Treussart, A. Clouqueur, C. Grossman, J.-F. Roch, Photon antibunching in the fluorescence of a single dye molecule embedded in a thin polymer film. *Opt. Lett.* **26**, 1504–1506 (2001)
54. F. Vargas, O. Hollricher, O. Marti, G. de Schaetzen, G. Tarrach, Influence of protective layers on the blinking of fluorescent single molecules observed by confocal microscopy and scanning near field optical microscopy. *J. Chem. Phys.* **117**, 866–871 (2002)
55. L.A. Deschenes, D.A.V. Bout, Single-molecule studies of heterogeneous dynamics in polymer melts near the glass transition. *Science* **292**, 255–258 (2001)
56. S.G. Lukishova, A.W. Schmid, A.J. McNamara, R.W. Boyd, C.R. Stroud, Room temperature single-photon source: single-dye molecule fluorescence in liquid crystal host. *IEEE J Sel. Top. Quant. Electron.* **9**, 1512–1518 (2003)

57. C. Kurtsiefer, S. Mayer, P. Zarda, H. Weinfurter, Stable solid-state source of single photons. *Phys. Rev. Lett.* **85**, 290 (2000)
58. R. Brouri, A. Beveratos, J.-P. Poizat, P. Grangier, Photon antibunching in the fluorescence of individual color centers in diamond. *Opt. Lett.* **25**, 1294–1296 (2000)
59. A. Beveratos, R. Brouri, T. Gacoin, J.-P. Poizat, P. Grangier, Nonclassical radiation from diamond nanocrystals. *Phys. Rev. A* **64**, 61802 (2001)
60. A. Beveratos, R. Brouri, T. Gacoin, A. Villing, J.-P. Poizat, P. Grangier, Single photon quantum cryptography. *Phys. Rev. Lett.* **89**, 187901 (2002)
61. P. Michler, A. Imamoğlu, M.D. Mason, P.J. Carson, G.F. Strouse, S.K. Buratto, Quantum correlation among photons from a single quantum dot at room temperature. *Nature* **406**, 968–970 (2000)
62. B. Lounis, H.A. Bechtel, D. Gerion, P. Alivisatos, W.E. Moerner, Photon antibunching in single CdSe/ZnS quantum dot fluorescence. *Chem. Phys. Lett.* **329**, 399–404 (2000)
63. G. Messin, J.P. Hermier, E. Giacobino, P. Desbiolles, M. Dahan, Bunching and antibunching in the fluorescence of semiconductor nanocrystals. *Opt. Lett.* **26**, 1891–1893 (2001)
64. J. Kim, O. Benson, H. Kan, Y. Yamamoto, A single-photon turnstile device. *Nature* **397**, 500–503 (1999)
65. P. Michler, A. Kiraz, C. Becher, W.V. Schoenfeld, P.M. Petroff, L. Zhang, E. Hu, A. Imamoglu, A quantum dot single-photon turnstile device. *Science* **290**, 2282–2285 (2000)
66. E. Moreau, I. Robert, J.M. Gerard, I. Abram, L. Manin, V. Thierry-Mieg, Single-mode solid-state single photon source based on isolated quantum dots in pillar microcavities. *Appl. Phys. Lett.* **79**, 2865–2867 (2001)
67. C. Santori, M. Pelton, G. Solomon, Y. Dale, Y. Yamamoto, Triggered single photons from a quantum dot. *Phys. Rev. Lett.* **86**, 1502–1505 (2001)
68. Z. Yuan, B.E. Kardynal, R.M. Stevenson, A.J. Shields, C.J. Lobo, K. Cooper, N.S. Beattie, D.A. Ritchie, M. Pepper, Electrically driven single-photon source. *Science* **295**, 102–105 (2002)
69. M.D. Eisaman, J. Fan, A. Migdall, S.V. Polyakov, Invited review article: single-photon sources and detectors. *Rev. Sci. Instrum.* **82**, 71101 (2011)
70. Y. Yamamoto, C. Santori, G. Solomon, J. Vuckovic, D. Fattal, E. Waks, E. Diamanti, Single photons for quantum information systems. *Prog. Inform.* **5** (2005)
71. P. Yao, V.S.C. Manga Rao, S. Hughes, On-chip single photon sources using planar photonic crystals and single quantum dots. *Laser Photonics Rev.* **4**, 499–516 (2010)
72. S. Hughes: Single-photon sources: dream and reality, in *Single-Photon Generation and Detection: Physics and Applications*, A. Migdall, S.V. Polyakov, J. Fan, J.C. Bienfang (Academic Press, 2013)
73. B. Lounis, M. Orrit, Single-photon sources. *Rep. Prog. Phys.* **68**, 1129–1179 (2005)
74. K.O. Greulich, E. Thiel, Single photon light sources. *Single Mol.* **2**, 5–12 (2001)
75. M. Oxborrow, A.G. Sinclair, Single-photon sources. *Contemp. Phys.* **46**, 173–206 (2005)
76. A. Kuhn, D. Ljunggren, Cavity-based single-photon sources. *Contemp. Phys.* **51**, 289–313 (2010)
77. S. Scheel, Single-photon sources—an introduction. *J. Mod. Opt.* **56**, 141 (2009)
78. M. De Vittorio, F. Pisanello, L. Martiradonna, A. Qualtieri, T. Stomeo, A. Bramati, R. Cingolani, Recent advances on single photon sources based on single colloidal nanocrystals. *Opto-Electron. Rev.* **18**, 1–9 (2010)
79. L.J. Bissell, Experimental realization of efficient, room temperature single-photon sources with definite circular and linear polarizations. Ph.D. thesis, University of Rochester, Rochester, NY (2011)
80. S.G. Lukishova, Single photon sources for secure quantum communication, in *Proceedings SPIE 9065*, paper 90650C (2013)
81. K.S. Grußmayer, D.-P. Herten, Photon antibunching in single molecule fluorescence spectroscopy, in *Advanced Photon Counting*. Springer Series on Fluorescence, ed. by P. Kapusta, M. Wahl, R. Erdmann, vol. 15 (Springer International Publishing, 2014), pp. 159–190

82. P. Grangier, B. Sanders, J. Vuckovic (eds): Focus on single photons on demand. *New J. Phys.* **6**, (2004)
83. C.J. Chunnillall, I.P. Degiovanni, S. Kück, I. Müller, A.G. Sinclair, Metrology of single-photon sources and detectors: a review. *Opt. Eng.* **53**, 081910–081910 (2014)
84. G.S. Buller, R.J. Collins, Single-photon generation and detection. *Meas. Sci. Technol.* **21**, 12002 (2010)
85. E. Neu, C. Becher, Diamond-based single-photon sources and their application in quantum key distribution, in *Quantum Information Processing with Diamond: Principles and Applications*, ed. by S. Prawer, I. Aharonovich (Elsevier, 2014)
86. C. Santori, D. Fattal, Y. Yamamoto, *Single-Photon Devices and Applications* (Wiley, 2010)
87. P. Michler, *Single Semiconductor Quantum Dots* (Springer, Berlin, Heidelberg, 2010)
88. A. Migdall, S. Polyakov, J. Fan, J. Bienfang, *Single-Photon Generation and Detection: Physics and Applications*, (Academic Press, 2013)
89. E.M. Purcell, Spontaneous emission at radio frequencies. *Phys. Rev.* **69**, 681 (1946)
90. L. Novotny, B. Hecht, *Principles of Nano-optics* (Cambridge University Press, 2006)
91. A.F. Koenderink, On the use of Purcell factors for plasmon antennas. *Opt. Lett.* **35**, 4208 (2010)
92. A. Aspect, P. Grangier, The first single-photon sources, in *Single-Photon Generation and Detection*, ed. by A. Migdall, S.V. Polyakov, J. Fan, J.C. Bienfang (Elsevier, 2013), pp. 315–350
93. S.A. Castelletto, R.E. Scholten, Heralded single photon sources: a route towards quantum communication technology and photon standards. *Eur. Phys. J.—Appl. Phys.* **41**, 181–194 (2008)
94. H. Huang, A. Dorn, V. Bulovic, M.G. Bawendi, Electrically driven light emission from single colloidal quantum dots at room temperature. *Appl. Phys. Lett.* **90**, 23110–23113 (2007)
95. S. Brovelli, W.K. Bae, C. Galland, U. Giovanella, F. Meinardi, V.I. Klimov, Dual-color electroluminescence from dot-in-bulk nanocrystals. *Nano Lett.* **14**, 486–494 (2014)
96. W.K. Bae, Y.-S. Park, J. Lim, D. Lee, L.A. Padilha, H. McDaniel, I. Robel, C. Lee, J.M. Pietryga, V.I. Klimov, Controlling the influence of Auger recombination on the performance of quantum-dot light-emitting diodes. *Nat. Commun.* **4**, 2661 (2013)
97. B.N. Pal, I. Robel, A. Mohite, R. Laocharoensuk, D.J. Werder, V.I. Klimov, High-sensitivity p–n junction photodiodes based on PbS nanocrystal quantum dots. *Adv. Funct. Mater.* **22**, 1741–1748 (2012)
98. B.N. Pal, Y. Ghosh, S. Brovelli, R. Laocharoensuk, V.I. Klimov, J.A. Hollingsworth, H. Htoon, “Giant” CdSe/CdS core/shell nanocrystal quantum dots as efficient electroluminescent materials: strong influence of shell thickness on light-emitting diode performance. *Nano Lett.* **12**, 331–336 (2012)
99. N. Mizuochi, T. Makino, H. Kato, D. Takeuchi, M. Ogura, H. Okushi, M. Nothaft, P. Neumann, A. Gali, F. Jelezko, J. Wrachtrup, S. Yamasaki, Electrically driven single-photon source at room temperature in diamond. *Nat. Photonics* **6**, 299–303 (2012)
100. J. Forneris, P. Traina, D.G. Monticone, G. Amato, L. Boarino, G. Brida, I.P. Degiovanni, E. Enrico, E. Moreva, V. Grilj, N. Skukan, M. Jakšić, M. Genovese, P. Oliviero, Electrical stimulation of non-classical photon emission from diamond color centers by means of sub-superficial graphitic electrodes. *Sci. Rep.* **5**, 15901 (2015)
101. M. Nothaft, S. Höhla, F. Jelezko, N. Frühauf, J. Pflaum, J. Wrachtrup, Electrically driven photon antibunching from a single molecule at room temperature. *Nat. Commun.* **3**, 628 (2012)
102. J.I. Gonzalez, T.-H. Lee, M.D. Barnes, Y. Antoku, R.M. Dickson, Quantum mechanical single-gold-nanocluster electroluminescent light source at room temperature. *Phys. Rev. Lett.* **93**, 147402 (2004)
103. V.I. Klimov, *Nanocrystal Quantum Dots*, 2nd edn. (CRC Press, 2010)
104. A.L. Rogach (ed.), *Semiconductor Nanocrystal Quantum Dots* (Springer, Vienna, 2008)

105. J.M. Pietryga, R.D. Schaller, D. Werder, M.H. Stewart, V.I. Klimov, J.A. Hollingsworth, Pushing the band gap envelope: mid-infrared emitting colloidal PbSe quantum dots. *J. Am. Chem. Soc.* **126**, 11752–11753 (2004)
106. H. Du, C. Chen, R. Krishnan, T.D. Krauss, J.M. Harbold, F.W. Wise, M.G. Thomas, J. Silcox, Optical properties of colloidal PbSe nanocrystals. *Nano Lett.* **2**, 1321–1324 (2002)
107. J. Heo, T. Zhu, C. Zhang, J. Xu, P. Bhattacharya, Electroluminescence from silicon-based photonic crystal microcavities with PbSe quantum dots. *Opt. Lett.* **35**, 547–549 (2010)
108. S.G. Lukishova, L.J. Bissell, V.M. Menon, N. Valappil, M.A. Hahn, C.M. Evans, B. Zimmerman, T.D. Krauss, C.R. Stroud, R.W. Boyd, Organic photonic bandgap microcavities doped with semiconductor nanocrystals for room-temperature on-demand single-photon sources. *J. Mod. Opt.* **56**, 167–174 (2009)
109. Z. Wu, Z. Mi, P. Bhattacharya, T. Zhu, J. Xu, Enhanced spontaneous emission at 1.55 μm from colloidal PbSe quantum dots in a Si photonic crystal microcavity. *Appl. Phys. Lett.* **90**, 171105–3 (2007)
110. J. Yang, J. Heo, T. Zhu, J. Xu, J. Topolancik, F. Vollmer, R. Ilic, P. Bhattacharya, Enhanced photoluminescence from embedded PbSe colloidal quantum dots in silicon-based random photonic crystal microcavities. *Appl. Phys. Lett.* **92**, 261110 (2008)
111. M.T. Rakher, R. Bose, C.W. Wong, K. Srinivasan, Spectroscopy of 1.55 μm PbS quantum dots on Si photonic crystal cavities with a fiber taper waveguide. *Appl. Phys. Lett.* **96**, 161108 (2010)
112. A.G. Pattantyus-Abraham, H. Qiao, J. Shan, K.A. Abel, T.-S. Wang, F.C.J.M. van Veggel, J.F. Young, Site-selective optical coupling of PbSe nanocrystals to Si-based photonic crystal microcavities. *Nano Lett.* **9**, 2849–2854 (2009)
113. C.A. Foell, E. Schelew, H. Qiao, K.A. Abel, S. Hughes, F.C.J.M. van Veggel, J.F. Young, Saturation behaviour of colloidal PbSe quantum dot exciton emission coupled into silicon photonic circuits. *Opt. Express* **20**, 10453–10469 (2012)
114. Y. Chen, J. Vela, H. Htoon, J.L. Casson, D.J. Werder, D.A. Bussian, V.I. Klimov, J.A. Hollingsworth, “Giant” multishell CdSe nanocrystal quantum dots with suppressed blinking. *J. Am. Chem. Soc.* **130**, 5026–5027 (2008)
115. O. Chen, J. Zhao, V.P. Chauhan, J. Cui, C. Wong, D.K. Harris, H. Wei, H.-S. Han, D. Fukumura, R.K. Jain, M.G. Bawendi, Compact high-quality CdSe–CdS core–shell nanocrystals with narrow emission linewidths and suppressed blinking. *Nat. Mater.* **12**, 445–451 (2013)
116. A.M. Keller, Y. Ghosh, M.S. DeVore, M.E. Phipps, M.H. Stewart, B.S. Wilson, D.S. Lidke, J.A. Hollingsworth, J.H. Werner, 3-Dimensional tracking of non-blinking “giant” quantum dots in live cells. *Adv. Funct. Mater.* **24**, 4796–4803 (2014)
117. B. Mahler, P. Spinicelli, S. Buil, X. Quelin, J.-P. Hermier, B. Dubertret, Towards non-blinking colloidal quantum dots. *Nat. Mater.* **7**, 659–664 (2008)
118. F. García-Santamaría, Y. Chen, J. Vela, R.D. Schaller, J.A. Hollingsworth, V.I. Klimov, Suppressed Auger recombination in “giant” nanocrystals boosts optical gain performance. *Nano Lett.* **9**, 3482–3488 (2009)
119. M. Manceau, S. Vezzoli, Q. Glorieux, F. Pisanello, E. Giacobino, L. Carbone, M. De Vittorio, A. Bramati, Effect of charging on CdSe/CdS dot-in-rods single-photon emission. *Phys. Rev. B* **90**, 35311 (2014)
120. F. Pisanello, G. Leménager, L. Martiradonna, L. Carbone, S. Vezzoli, P. Desfonds, P.D. Cozzoli, J.-P. Hermier, E. Giacobino, R. Cingolani, M. De Vittorio, A. Bramati, Non-blinking single-photon generation with anisotropic colloidal nanocrystals: towards room-temperature, efficient, colloidal quantum sources. *Adv. Mater.* **25**, 1974–1980 (2013)
121. F. Pisanello, L. Martiradonna, G. Leménager, P. Spinicelli, A. Fiore, L. Manna, J.-P. Hermier, R. Cingolani, E. Giacobino, M. De Vittorio, A. Bramati, Room temperature-dipolelike single photon source with a colloidal dot-in-rod. *Appl. Phys. Lett.* **96**, 33101 (2010)

122. T. Du, J. Schneider, A.K. Srivastava, A.S. Susha, V.G. Chigrinov, H.S. Kwok, A.L. Rogach, Combination of photoinduced alignment and self-assembly to realize polarized emission from ordered semiconductor nanorods. *ACS Nano* **9**, 11049–11055 (2015)
123. A.M. Zaitsev, *Optical Properties of Diamond: A Data Handbook* (Springer Science & Business Media, 2001)
124. R. Mildren, J. Rabeau, *Optical Engineering of Diamond* (Wiley, 2013)
125. S. Prawer, I. Aharonovich, *Quantum Information Processing with Diamond: Principles and Applications* (Elsevier, 2014)
126. I. Aharonovich, A.D. Greentree, S. Prawer, Diamond photonics. *Nat. Photonics* **5**, 397–405 (2011)
127. I. Aharonovich, S. Castelletto, D.A. Simpson, C.-H. Su, A.D. Greentree, S. Prawer, Diamond-based single-photon emitters. *Rep. Prog. Phys.* **74**, 76501 (2011)
128. A. Gruber, A. Drabenstedt, C. Tietz, L. Fleury, J. Wrachtrup, C. von Borczyskowski, Scanning confocal optical microscopy and magnetic resonance on single defect centers. *Science* **276**, 2012–2014 (1997)
129. B. Naydenov, R. Kolesov, A. Batalov, J. Meijer, S. Pezzagna, D. Rogalla, F. Jelezko, J. Wrachtrup, Engineering single photon emitters by ion implantation in diamond. *Appl. Phys. Lett.* **95**, 181109 (2009)
130. A.T. Collins, M.F. Thomaz, M.I.B. Jorge, Luminescence decay time of the 1.945 eV centre in type Ib diamond. *J. Phys. C Solid State Phys.* **16**, 2177 (1983)
131. A. Beveratos, S. Kühn, R. Brouri, T. Gacoin, J.P. Poizat, P. Grangier, Room temperature stable single-photon source. *Eur. Phys. J.—At. Mol. Opt. Plasma Phys.* **18**, 191–196 (2002)
132. H. Bernien, L. Childress, L. Robledo, M. Markham, D. Twitchen, R. Hanson, Two-photon quantum interference from separate nitrogen vacancy centers in diamond. *Phys. Rev. Lett.* **108**, 43604 (2012)
133. A. Mohtashami, A.F. Koenderink, Suitability of nanodiamond nitrogen–vacancy centers for spontaneous emission control experiments. *New J. Phys.* **15**, 43017 (2013)
134. J.-H. Hsu, W.-D. Su, K.-L. Yang, Y.-K. Tzeng, H.-C. Chang, Nonblinking green emission from single H3 color centers in nanodiamonds. *Appl. Phys. Lett.* **98**, 193116 (2011)
135. D.A. Simpson, E. Ampem-Lassen, B.C. Gibson, S. Trpkovski, F.M. Hossain, S.T. Huntington, A.D. Greentree, L.C.L. Hollenberg, S. Prawer, A highly efficient two level diamond based single photon source. *Appl. Phys. Lett.* **94**, 203107–203107-3 (2009)
136. J.M. Smith, F. Grazioso, B.R. Patton, P.R. Dolan, M.L. Markham, D.J. Twitchen, Optical properties of a single-colour centre in diamond with a green zero-phonon line. *New J. Phys.* **13**, 45005 (2011)
137. C. Wang, C. Kurtsiefer, H. Weinfurter, B. Burchard, Single photon emission from SiV centres in diamond produced by ion implantation. *J. Phys. B: At. Mol. Opt. Phys.* **39**, 37–41 (2006)
138. E. Neu, D. Steinmetz, J. Riedrich-Möller, S. Gsell, M. Fischer, M. Schreck, C. Becher, Single photon emission from silicon-vacancy colour centres in chemical vapour deposition nano-diamonds on iridium. *New J. Phys.* **13**, 25012 (2011)
139. E. Neu, M. Agio, C. Becher, Photophysics of single silicon vacancy centers in diamond: implications for single photon emission. *Opt. Express* **20**, 19956–19971 (2012)
140. E. Neu, M. Fischer, S. Gsell, M. Schreck, C. Becher, Fluorescence and polarization spectroscopy of single silicon vacancy centers in heteroepitaxial nanodiamonds on iridium. *Phys. Rev. B* **84**, 205211 (2011)
141. I. Aharonovich, S. Castelletto, B.C. Johnson, J.C. McCallum, D.A. Simpson, A.D. Greentree, S. Prawer, Chromium single-photon emitters in diamond fabricated by ion implantation. *Phys. Rev. B* **81**, 121201 (2010)
142. I. Aharonovich, S. Castelletto, D.A. Simpson, A. Stacey, J. McCallum, A.D. Greentree, S. Prawer, Two-level ultrabright single photon emission from diamond nanocrystals. *Nano Lett.* **9**, 3191–3195 (2009)
143. S. Castelletto, A. Boretti, Radiative and nonradiative decay rates in chromium-related centers in nanodiamonds. *Opt. Lett.* **36**, 4224–4226 (2011)

144. D. Steinmetz, E. Neu, J. Meijer, W. Bolse, C. Becher, Single photon emitters based on Ni/Si related defects in single crystalline diamond. *Appl. Phys. B* **102**, 451–458 (2011)
145. I. Aharonovich, C. Zhou, A. Stacey, J. Orwa, S. Castelletto, D. Simpson, A.D. Greentree, F. Treussart, J.-F. Roch, S. Prawer, Enhanced single-photon emission in the near infrared from a diamond color center. *Phys. Rev. B* **79**, 235316 (2009)
146. T. Gaebel, I. Popa, A. Gruber, M. Domhan, F. Jelezko, J. Wrachtrup, Stable single-photon source in the near infrared. *New J. Phys.* **6**, 98–98 (2004)
147. E. Wu, V. Jacques, H. Zeng, P. Grangier, F. Treussart, J.-F. Roch, Narrow-band single-photon emission in the near infrared for quantum key distribution. *Opt. Express* **14**, 1296–1303 (2006)
148. J.R. Rabeau, Y.L. Chin, S. Prawer, F. Jelezko, T. Gaebel, J. Wrachtrup, Fabrication of single nickel-nitrogen defects in diamond by chemical vapor deposition. *Appl. Phys. Lett.* **86**, 131926 (2005)
149. E. Wu, J. Rabeau, G. Roger, F. Treussart, H. Zeng, P. Grangier, S. Prawer, J. Roch, Room temperature triggered single-photon source in the near infrared. *New J. Phys.* **9** (2007)
150. T. Iwasaki, F. Ishibashi, Y. Miyamoto, Y. Doi, S. Kobayashi, T. Miyazaki, K. Tahara, K.D. Jahnke, L.J. Rogers, B. Naydenov, F. Jelezko, S. Yamasaki, S. Nagamachi, T. Inubushi, N. Mizuuchi, M. Hatano, Germanium-vacancy single color centers in diamond. *Sci. Rep.* **5**, 12882 (2015)
151. H.-Q. Zhao, M. Fujiwara, S. Takeuchi, Suppression of fluorescence phonon sideband from nitrogen vacancy centers in diamond nanocrystals by substrate effect. *Opt. Express* **20**, 15628 (2012)
152. M. Leifgen, T. Schröder, F. Gädeke, R. Riemann, V. Métillon, E. Neu, C. Hepp, C. Arend, C. Becher, K. Lauritsen, O. Benson, Evaluation of nitrogen- and silicon-vacancy defect centres as single photon sources in quantum key distribution. *New J. Phys.* **16**, 23021 (2014)
153. A.D. Greentree, I. Aharonovich, S. Castelletto, M.W. Doherty, L.P. McGuinness, D.A. Simpson, 21st-century applications of nanodiamonds. *Opt. Photonics News* **21**, 20–25 (2010)
154. J.R. Rabeau, S.T. Huntington, A.D. Greentree, S. Prawer, Diamond chemical-vapor deposition on optical fibers for fluorescence waveguiding. *Appl. Phys. Lett.* **86**, 134104 (2005)
155. A.W. Schell, G. Kewes, T. Schröder, J. Wolters, T. Aichele, O. Benson, A scanning probe-based pick-and-place procedure for assembly of integrated quantum optical hybrid devices. *Rev. Sci. Instrum.* **82**, 73709 (2011)
156. T. Schröder, A.W. Schell, G. Kewes, T. Aichele, O. Benson, Fiber-integrated diamond-based single photon source. *Nano Lett.* **11**, 198–202 (2011)
157. T. Gaebel, C. Bradac, J. Chen, J.M. Say, L. Brown, P. Hemmer, J.R. Rabeau, Size-reduction of nanodiamonds via air oxidation. *Diam. Relat. Mater.* **21**, 28–32 (2012)
158. I.I. Vlasov, A.A. Shiryaev, T. Rendler, S. Steinert, S.-Y. Lee, D. Antonov, M. Vörös, F. Jelezko, A.V. Fisenko, L.F. Semjonova, J. Biskupek, U. Kaiser, O.I. Lebedev, I. Sildos, P.R. Hemmer, V.I. Konov, A. Gali, J. Wrachtrup, Molecular-sized fluorescent nanodiamonds. *Nat. Nanotechnol.* **9**, 54–58 (2014)
159. I.I. Vlasov, O. Shenderova, S. Turner, O.I. Lebedev, A.A. Basov, I. Sildos, M. Rähn, A.A. Shiryaev, G. Van Tendeloo, Nitrogen and luminescent nitrogen-vacancy defects in detonation nanodiamond. *Small* **6**, 687–694 (2010)
160. C. Bradac, T. Gaebel, N. Naidoo, M.J. Sellars, J. Twamley, L.J. Brown, A.S. Barnard, T. Plakhotnik, A.V. Zvyagin, J.R. Rabeau, Observation and control of blinking nitrogen-vacancy centres in discrete nanodiamonds. *Nat. Nanotechnol.* **5**, 345–349 (2010)
161. A. Wolcott, T. Schiros, M.E. Trusheim, E.H. Chen, D. Nordlund, R.E. Diaz, O. Gaathon, D. Englund, J.S. Owen, Surface structure of aerobically oxidized diamond nanocrystals. *J. Phys. Chem. C* **118**, 26695–26702 (2014)
162. R. Kumar, M. Nyk, T.Y. Ohulchanskyy, C.A. Flask, P.N. Prasad, Combined optical and MR bioimaging using rare earth ion doped NaYF₄ nanocrystals. *Adv. Funct. Mater.* **19**, 853–859 (2009)

163. S. Schietinger, T. Aichele, H.-Q. Wang, T. Nann, O. Benson, Plasmon-enhanced upconversion in single $\text{NaYF}_4\text{:Yb}^{3+}\text{/Er}^{3+}$ codoped nanocrystals. *Nano Lett.* **10**, 134–138 (2010)
164. S. Schietinger, L.S. de Menezes, B. Lauritzen, O. Benson, Observation of size dependence in multicolor upconversion in single Yb^{3+} , Er^{3+} codoped NaYF_4 nanocrystals. *Nano Lett.* **9**, 2477–2481 (2009)
165. S. Schietinger, Investigation, manipulation, and coupling of single nanoscopic and quantum emitters. Ph.D. thesis (Humboldt University, Berlin, 2012). <http://edoc.hu-berlin.de/dissertationen/schietinger-stefan-2012-03-01/PDF/schietinger.pdf>. Accessed 6 Mar 2016
166. S. Babu, J.-H. Cho, J.M. Dowding, E. Heckert, C. Komanski, S. Das, J. Colon, C.H. Baker, M. Bass, W.T. Self, S. Seal, Multicolored redox active upconverter cerium oxide nanoparticle for bio-imaging and therapeutics. *Chem. Commun.* **46**, 6915–6917 (2010)
167. R. Kolesov, K. Xia, R. Reuter, R. Stöhr, A. Zappe, J. Meijer, P.R. Hemmer, J. Wrachtrup, Optical detection of a single rare-earth ion in a crystal. *Nat. Commun.* **3**, 1029 (2012)
168. R. Kolesov, K. Xia, R. Reuter, R. Stöhr, P. Siyushev, T. Inal, M. Jamali, J. Wrachtrup, J. Meijer, P.R. Hemmer, Optical detection of single rare-earth species in a crystal. Presented at Photonics West, 2–7 Feb 2013, San Francisco, CA
169. Y. Ying, Single instance rare-earth quantum computing with single Ce-ion read out. Presented at Laser Physics Workshop, Seminar 7. Quantum Information Science, p. 22, CD (2012)
170. M. Barnes, A. Mehta, T. Thundat, R. Bhargava, A. Bartko, L. Peyser, R.M. Dickson, Probing single ion luminescence in rare-earth doped nanocrystals. Presented at Laser Applications to Chemical and Environmental Analysis, p. ThA3 (Optical Society of America, 2002)
171. M.D. Barnes, A. Mehta, T. Thundat, R.N. Bhargava, V. Chhabra, B. Kulkarni, On-off blinking and multiple bright states of single europium ions in $\text{Eu}^{3+}\text{:Y}_2\text{O}_3$ nanocrystals. *J. Phys. Chem. B.* **104**, 6099–6102 (2000)
172. A.P. Bartko, L.A. Peyser, R.M. Dickson, A. Mehta, T. Thundat, R. Bhargava, M.D. Barnes, Observation of dipolar emission patterns from isolated $\text{Eu}^{3+}\text{:Y}_2\text{O}_3$ doped nanocrystals: new evidence for single ion luminescence. *Chem. Phys. Lett.* **358**, 459–465 (2002)
173. A. Mehta, T. Thundat, M.D. Barnes, V. Chhabra, R. Bhargava, A.P. Bartko, R.M. Dickson, Size-correlated spectroscopy and imaging of rare-earth-doped nanocrystals. *Appl. Opt.* **42**, 2132–2139 (2003)
174. E. Eichhammer, T. Utikal, S. Götzinger, V. Sandoghdar, Spectroscopic detection of single Pr^{3+} ions on the ${}^3\text{H}_4\text{--}{}^1\text{D}_2$ transition. *New J. Phys.* **17**, 83018 (2015)
175. C. Yin, M. Rancic, G.G. de Boo, N. Stavrias, J.C. McCallum, M.J. Sellars, S. Rogge, Optical addressing of an individual erbium ion in silicon. *Nature* **497**, 91–94 (2013)
176. T. Utikal, E. Eichhammer, L. Petersen, A. Renn, S. Götzinger, V. Sandoghdar, Spectroscopic detection and state preparation of a single praseodymium ion in a crystal. *Nat. Commun.* **5**, 3627 (2014)
177. P. Siyushev, K. Xia, R. Reuter, M. Jamali, N. Zhao, N. Yang, C. Duan, N. Kukharchyk, A.D. Wieck, R. Kolesov, J. Wrachtrup, Coherent properties of single rare-earth spin qubits. *Nat. Commun.* **5**, 3895 (2014)
178. R. Kolesov, R. Reuter, K. Xia, R. Stöhr, A. Zappe, J. Wrachtrup, Super-resolution upconversion microscopy of praseodymium-doped yttrium aluminum garnet nanoparticles. *Phys. Rev. B.* **84**, 153413 (2011)
179. S.K. Gayen, B.Q. Xie, Y.M. Cheung, Two-photon excitation of the lowest $4\text{f}2\text{--}4\text{f}5\text{d}$ near-ultraviolet transitions in $\text{Pr}^{3+}\text{:Y}_3\text{Al}_5\text{O}_{12}$. *Phys. Rev. B.* **45**, 20–28 (1992)
180. J. Pejchal, M. Nikl, E. Mihokova et al., Pr^{3+} -doped complex oxide single crystal scintillators. *J. Phys. D Appl. Phys.* **42**, 055117 (2009)
181. M.J. Weber, Nonradiative decay from 5d states of rare-earths in crystals. *Solid State Commun.* **12**, 741–744 (1973)

182. R. Kolesov, K. Xia, R. Reuter, M. Jamali, R. Stöhr, T. Inal, P. Siyushev, J. Wrachtrup, Mapping spin coherence of a single rare-earth ion in a crystal onto a single photon polarization state. *Phys. Rev. Lett.* **111**, 120502 (2013)
183. R.R. Jacobs, W.F. Krupke, M.J. Weber, Measurement of excited-state-absorption loss for Ce³⁺ in Y₃Al₅O₁₂ and implications for tunable 5d-4f rare-earth lasers. *Appl. Phys. Lett.* **33** (5), 410–412 (1978)
184. S. Castelletto, B.C. Johnson, V. Ivády, N. Stavrias, T. Umeda, A. Gali, T. Ohshima, A silicon carbide room-temperature single-photon source. *Nat. Mater.* **13**, 151–156 (2014)
185. F. Fuchs, B. Stender, M. Trupke, D. Simin, J. Pflaum, V. Dyakonov, G.V. Astakhov, Engineering near-infrared single-photon emitters with optically active spins in ultrapure silicon carbide. *Nat. Commun.* **6**, 7578 (2015)
186. F. Fuchs, V.A. Soltamov, S. Väth, P.G. Baranov, E.N. Mokhov, G.V. Astakhov, V. Dyakonov, Silicon carbide light-emitting diode as a prospective room temperature source for single photons. *Sci. Rep.* **3**, 1637 (2013)
187. A. Lohrmann, N. Iwamoto, Z. Bodrog, S. Castelletto, T. Ohshima, T.J. Karle, A. Gali, S. Prawer, J.C. McCallum, B.C. Johnson, Single-photon emitting diode in silicon carbide. *Nat. Commun.* **6**, 7783 (2015)
188. A. Lohrmann, B.C. Johnson, J.C. McCallum, S. Castelletto, A review on single photon sources in silicon carbide. *Rep. Prog. Phys.* **80**, 034502 (2017)
189. A.J. Morfa, B.C. Gibson, M. Karg, T.J. Karle, A.D. Greentree, P. Mulvaney, S. Tomljenovic-Hanic, Single-photon emission and quantum characterization of zinc oxide defects. *Nano Lett.* **12**, 949–954 (2012)
190. S. Choi, B.C. Johnson, S. Castelletto, C. Ton-That, M.R. Phillips, I. Aharonovich, Single photon emission from ZnO nanoparticles. *Appl. Phys. Lett.* **104**, 261101 (2014)
191. N.R. Jungwirth, H.-S. Chang, M. Jiang, G.D. Fuchs, Polarization spectroscopy of defect-based single photon sources in ZnO. *ACS Nano* **10**, 1210–1215 (2016)
192. S. Choi, A.M. Berhane, A. Gentle, C. Ton-That, M.R. Phillips, I. Aharonovich, Electroluminescence from localized defects in zinc oxide: toward electrically driven single photon sources at room temperature. *ACS Appl. Mater. Interfaces.* **7**, 5619–5623 (2015)
193. X. He, N.F. Hartmann, X. Ma, Y. Kim, R. Ihly, J.L. Blackburn, W. Gao, J. Kono, Y. Yomogida, A. Hirano, T. Tanaka, H. Kataura, H. Htoon, S.K. Doorn, Tunable room-temperature single-photon emission at telecom wavelengths from sp³ defects in carbon nanotubes. *Nat. Photonics* **11**, 577–582 (2017)
194. T.T. Tran, K. Bray, M.J. Ford, M. Toth, I. Aharonovich, Quantum emission from hexagonal boron nitride monolayers. *Nat. Nanotechnol.* **11**, 37–41 (2016)
195. F.T. Rabouw, N.M.B. Cogan, A.C. Berends, W. van der Stam, D. Vanmaekelbergh, A.F. Koenderink, T.D. Krauss, C. de M. Donega, Non-blinking single-photon emitters in silica. *Sci. Rep.* **6**, 21187 (2016)
196. J. Wrachtrup, 2D materials: single photons at room temperature. *Nat. Nanotechnol.* **11**, 7–8 (2016)
197. I. Aharonovich, D. Englund, M. Toth, Solid-state single-photon emitters. *Nat. Photonics* **10**, 631–641 (2016)
198. O. Benson, Assembly of hybrid photonic architectures from nanophotonic constituents. *Nature* **480**, 193–199 (2011)
199. X. Brokmann, E. Giacobino, M. Dahan, J.P. Hermier, Highly efficient triggered emission of single photons by colloidal CdSe/ZnS nanocrystals. *Appl. Phys. Lett.* **85**, 712–714 (2004)
200. T. Schröder, F. Gädke, M.J. Banholzer, O. Benson, Ultrabright and efficient single-photon generation based on nitrogen-vacancy centres in nanodiamonds on a solid immersion lens. *New J. Phys.* **13**, 55017 (2011)
201. L. Marseglia, J.P. Hadden, A.C. Stanley-Clarke, J.P. Harrison, B. Patton, Y.-L.D. Ho, B. Naydenov, F. Jelezko, J. Meijer, P.R. Dolan, J.M. Smith, J.G. Rarity, J.L. O'Brien, Nanofabricated solid immersion lenses registered to single emitters in diamond. *Appl. Phys. Lett.* **98**, 133107 (2011)

202. K.G. Lee, X.W. Chen, H. Eghlidi, P. Kukura, R. Lettow, A. Renn, V. Sandoghdar, S. Götzinger, A planar dielectric antenna for directional single-photon emission and near-unity collection efficiency. *Nat Photon.* **5**, 166–169 (2011)
203. X.-W. Chen, S. Götzinger, V. Sandoghdar, 99% efficiency in collecting photons from a single emitter. *Opt. Lett.* **36**, 3545–3547 (2011)
204. X.-L. Chu, T.J.K. Brenner, X.-W. Chen, Y. Ghosh, J.A. Hollingsworth, V. Sandoghdar, S. Götzinger, Experimental realization of an optical antenna designed for collecting 99% of photons from a quantum emitter. *Optica* **1**, 203–208 (2014)
205. S. Haroche, J.-M. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, 2006)
206. G. Grynberg, A. Aspect, C. Fabre, *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light* (Cambridge University Press, 2010)
207. P.R. Berman, *Cavity Quantum Electrodynamics* (Academic Press, 1994)
208. R.K. Chang, A.J. Campillo, *Optical Processes in Microcavities* (World Scientific, 1996)
209. H. Yokoyama, K. Ujihara, *Spontaneous Emission and Laser Oscillation in Microcavities* (CRC Press, 1995)
210. H. Benisty, J.-M. Gerard, R. Houdre, J. Rarity, C. Weisbuch, *Confined Photon Systems: Fundamentals and Applications* (Springer, Berlin, Heidelberg, 1999)
211. K.J. Vahala, Optical microcavities. *Nature* **424**, 839–846 (2003)
212. A. Badolato, K. Hennessy, M. Atatüre, J. Dreiser, E. Hu, P.M. Petroff, A. Imamoglu, Deterministic coupling of single quantum dots to single nanocavity modes. *Science* **308**, 1158–1161 (2005)
213. K. Hennessy, A. Badolato, M. Winger, D. Gerace, M. Atatüre, S. Gulde, S. Falt, E.L. Hu, A. Imamoglu, Quantum nature of a strongly coupled single quantum dot-cavity system. *Nature* **445**, 896–899 (2007)
214. S. Strauf, F. Jahnke, Single quantum dot nanolaser. *Laser Photonics Rev.* **5**, 607–633 (2011)
215. S. Noda, M. Fujita, T. Asano, Spontaneous-emission control by photonic crystals and nanocavities. *Nat Photonics* **1**, 449–458 (2007)
216. S.G. Lukishova, J.M. Winkler, L.J. Bissell, D. Mihaylova, A.C. Liapis, Z. Shi, D. Goldberg, V.M. Menon, R.W. Boyd, G. Chen, P. Prasad, Room-temperature single-photon sources based on nanocrystal fluorescence in photonic/plasmonic nanostructures, in *Proceedings of SPIE, Emerging Technologies in Security and Defense II; and Quantum-Physics-based Information Security III*, **9254**, paper 9254-05, Oct 2014
217. A.C. Liapis, B. Gao, M.R. Siddiqui, Z. Shi, R.W. Boyd, On-chip spectroscopy with thermally tuned high-Q photonic crystal cavities. *Appl. Phys. Lett.* **108**, 21105 (2016)
218. Y. Lai, S. Pirotta, G. Urbinati, D. Gerace, M. Minkov, V. Savona, A. Badolato, M. Galli, Genetically designed L3 photonic crystal nanocavities with measured quality factor exceeding one million. *Appl. Phys. Lett.* **104**, 241101 (2014)
219. H.Y. Ryu, M. Notomi, Enhancement of spontaneous emission from the resonant modes of a photonic crystal slab single-defect cavity. *Opt. Lett.* **28**, 2390–2392 (2003)
220. L. Sanchis, M.J. Cryan, J. Pozo, I.J. Craddock, J.G. Rarity, Ultrahigh Purcell factor in photonic crystal slab microcavities. *Phys. Rev. B* **76**, 45118 (2007)
221. D. Press, S. Gotzinger, S. Reitzenstein, C. Hofmann, A. Löffler, M. Kamp, A. Forchel, Y. Yamamoto, Photon Antibunching from a single quantum-dot-microcavity system in the strong coupling regime. *Phys. Rev. Lett.* **98**, 117402–117404 (2007)
222. M. Lončar, A. Faraon, Quantum photonic networks in diamond. *MRS Bull.* **38**, 144–148 (2013)
223. A. Faraon, C. Santori, Z. Huang, V.M. Acosta, R.G. Beausoleil, Coupling of nitrogen-vacancy centers to photonic crystal cavities in monocrystalline diamond. *Phys. Rev. Lett.* **109**, 33604 (2012)
224. B.J.M. Hausmann, B.J. Shields, Q. Quan, Y. Chu, N.P. de Leon, R. Evans, M.J. Burek, A.S. Zibrov, M. Markham, D.J. Twitchen, H. Park, M.D. Lukin, M. Lončar, Coupling of NV centers to photonic crystal nanobeams in diamond. *Nano Lett.* **13**, 5791–5796 (2013)

225. Q. Quan, M. Loncar, Deterministic design of wavelength scale, ultra-high Q photonic crystal nanobeam cavities. *Opt. Express* **19**, 18529–18542 (2011)
226. L. Li, T. Schröder, E.H. Chen, M. Walsh, I. Bayn, J. Goldstein, O. Gaathon, M.E. Trusheim, M. Lu, J. Mower, M. Cotlet, M.L. Markham, D.J. Twitchen, D. Englund, Coherent spin control of a nanocavity-enhanced qubit in diamond. *Nat. Commun.* **6**, 6173 (2015)
227. L. Li, E.H. Chen, J. Zheng, S.L. Mouradian, F. Dolde, T. Schröder, S. Karaveli, M.L. Markham, D.J. Twitchen, D. Englund, Efficient photon collection from a nitrogen vacancy center in a circular bullseye grating. *Nano Lett.* **15**, 1493–1497 (2015)
228. L. Li, T. Schröder, E.H. Chen, H. Bakhrus, D. Englund, One-dimensional photonic crystal cavities in single-crystal diamond. *Photonics Nanostruct.—Fundam. Appl.* **15**, 130–136 (2015)
229. J. Riedrich-Möller, C. Arend, C. Pauly, F. Mücklich, M. Fischer, S. Gsell, M. Schreck, C. Becher, Deterministic coupling of a single silicon-vacancy color center to a photonic crystal cavity in diamond. *Nano Lett.* **14**, 5281–5287 (2014)
230. J. Riedrich-Möller, L. Kipfstuhl, C. Hepp, E. Neu, C. Pauly, F. Mücklich, A. Baur, M. Wandt, S. Wolff, M. Fischer, S. Gsell, M. Schreck, C. Becher, One- and two-dimensional photonic crystal microcavities in single crystal diamond. *Nat. Nanotechnol.* **7**, 69–74 (2012)
231. M.J. Burek, N.P. de Leon, B.J. Shields, B.J.M. Hausmann, Y. Chu, Q. Quan, A.S. Zibrov, H. Park, M.D. Lukin, M. Lončar, Free-standing mechanical and photonic nanostructures in single-crystal diamond. *Nano Lett.* **12**, 6084–6089 (2012)
232. M.J. Burek, Y. Chu, M.S.Z. Liddy, P. Patel, J. Rochman, S. Meesala, W. Hong, Q. Quan, M.D. Lukin, M. Lončar, High quality-factor optical nanocavities in bulk single-crystal diamond. *Nat. Commun.* **5**, 5718 (2014)
233. B.J.M. Hausmann, I.B. Bulu, P.B. Deotare, M. McCutcheon, V. Venkataraman, M.L. Markham, D.J. Twitchen, M. Lončar, Integrated high-quality factor optical resonators in diamond. *Nano Lett.* **13**, 1898–1902 (2013)
234. C.F. Wang, R. Hanson, D.D. Awschalom, E.L. Hu, T. Feygelson, J. Yang, J.E. Butler, Fabrication and characterization of two-dimensional photonic crystal microcavities in nanocrystalline diamond. *Appl. Phys. Lett.* **91**, 201112 (2007)
235. T. Zhong, J. Rochman, J.M. Kindem, E. Miyazono, A. Faraon, High quality factor nanophotonic resonators in bulk rare-earth doped crystals. *Opt. Express* **24**, 536–544 (2016)
236. T. Zhong, J.M. Kindem, E. Miyazono, A. Faraon, Nanophotonic coherent light-matter interfaces based on rare-earth-doped crystals. *Nat. Commun.* **6**, 8206 (2015)
237. E. Miyazono, T. Zhong, I. Craiciu, J.M. Kindem, A. Faraon, Coupling of erbium dopants to yttrium orthosilicate photonic crystal cavities for on-chip optical quantum memories. *Appl. Phys. Lett.* **108**, 11111 (2016)
238. J. Wolters, A.W. Schell, G. Kewes, N. Nüssle, M. Schoengen, H. Dösscher, T. Hannappel, B. Löchel, M. Barth, O. Benson, Enhancement of the zero phonon line emission from a single nitrogen vacancy center in a nanodiamond via coupling to a photonic crystal cavity. *Appl. Phys. Lett.* **97**, 141108 (2010)
239. T. van der Sar, J. Hagemeier, W. Pfaff, E.C. Heeres, S.M. Thon, H. Kim, P.M. Petroff, T.H. Oosterkamp, D. Bouwmeester, R. Hanson, Deterministic nanoassembly of a coupled quantum emitter–photonic crystal cavity system. *Appl. Phys. Lett.* **98**, 193103 (2011)
240. D. Englund, B. Shields, K. Rivoire, F. Hatami, J. Vučković, H. Park, M.D. Lukin, Deterministic coupling of a single nitrogen vacancy center to a photonic crystal cavity. *Nano Lett.* **10**, 3922–3926 (2010)
241. I. Fushman, D. Englund, J. Vučković, Coupling of PbS quantum dots to photonic crystal cavities at room temperature. *Appl. Phys. Lett.* **87**, 241102–241103 (2005)
242. R. Bose, X. Yang, R. Chatterjee, J. Gao, C.W. Wong, Weak coupling interactions of colloidal lead sulphide nanocrystals with silicon photonic crystal nanocavities near $1.55\text{ }\mu\text{m}$ at room temperature. *Appl. Phys. Lett.* **90**, 111117–3 (2007)
243. D.F. Dorfner, T. Hürlimann, G. Abstreiter, J.J. Finley, Optical characterization of silicon on insulator photonic crystal nanocavities infiltrated with colloidal PbS quantum dots. *Appl. Phys. Lett.* **91**, 233111 (2007)

244. C.B. Poitras, M. Lipson, H. Du, M.A. Hahn, T.D. Krauss, Photoluminescence enhancement of colloidal quantum dots embedded in a monolithic microcavity. *Appl. Phys. Lett.* **82**, 4032–4034 (2003)
245. N. Valappil, M. Luberto, V.M. Menon, I. Zeylikovich, T.K. Gayen, J. Franco, B.B. Das, R.R. Alfano, Solution processed microcavity structures with embedded quantum dots. *Photonics Nanostruct—Fundam. Appl.* **5**, 184–188 (2007)
246. D. Goldberg, V.M. Menon, Enhanced amplified spontaneous emission from colloidal quantum dots in all-dielectric monolithic microcavities. *Appl. Phys. Lett.* **102**, 81119–81119-4 (2013)
247. S.G. Lukishova, J.M. Winkler, D. Mihaylova, A. Liapis, L.J. Bissell, David Goldberg, V.M. Menon, Z. Shi, R.W. Boyd, G. Chen, P. Prasad, Nanocrystal fluorescence in photonic bandgap microcavities and plasmonic nanoantennas. *J. Phys. Conf. Ser.* **594**, 12005 (2015)
248. L. Bissell, D. Goldberg, S.G. Lukishova, V.M. Menon, Quantum dot single-photon source in a Bragg reflector microcavity with a defect layer. Presented at Frontiers in Optics 2012/Laser Science XXVIII, p. LTh1I.2, 14–18 Oct, Rochester, NY (2012)
249. L. Martiradonna, M. De Vittorio, L. Troisi, M.T. Todaro, M. Mazzeo, T. Stomeo, M. Annì, R. Cingolani, G. Gigli, Fabrication of hybrid organic–inorganic vertical microcavities through imprint technology. *Microelectron. Eng.* **78–79**, 593–597 (2005)
250. L. Martiradonna, L. Carbone, M.D. Giorgi, L. Manna, G. Gigli, R. Cingolani, M.D. Vittorio, High Q-factor colloidal nanocrystal-based vertical microcavity by hot embossing technology. *Appl. Phys. Lett.* **88**, 181108 (2006)
251. A. Qualtieri, G. Morello, P. Spinicelli, M.T. Todaro, T. Stomeo, L. Martiradonna, M.D. Giorgi, X. Quelin, S. Buil, A. Bramati, J.P. Hermier, R. Cingolani, M.D. Vittorio, Nonclassical emission from single colloidal nanocrystals in a microcavity: a route towards room temperature single photon sources. *New J. Phys.* **11**, 33025 (2009)
252. A. Qualtieri, G. Morello, P. Spinicelli, M.T. Todaro, T. Stomeo, L. Martiradonna, M. De Giorni, X. Quelin, S. Buil, A. Bramati, J.P. Hermier, R. Cingolani, M. De Vittorio, Room temperature single-photon sources based on single colloidal nanocrystals in microcavities. *Superlattices Microstruct.* **47**, 187–191 (2010)
253. M. Kahl, T. Thomay, V. Kohnle, K. Beha, J. Merlein, M. Hagner, A. Halm, J. Ziegler, T. Nann, Y. Fedutik, U. Woggon, M. Artemyev, F. Perez-Willard, A. Leitenstorfer, R. Bratschitsch, Colloidal quantum dots in all-dielectric high-Q pillar microcavities. *Nano Lett.* **7**, 2897–2900 (2007)
254. T. Thomay, T. Hanke, M. Tomas, F. Sotier, K. Beha, V. Knittel, M. Kahl, K.M. Whitaker, D.R. Gamelin, A. Leitenstorfer, R. Bratschitsch, Colloidal ZnO quantum dots in ultraviolet pillar microcavities. *Opt. Express* **16**, 9791–9794 (2008)
255. J. Jasieniak, C. Sada, A. Chiasera, M. Ferrari, A. Martucci, P. Mulvaney, Sol–gel based vertical optical microcavities with quantum dot defect layers. *Adv. Funct. Mater.* **18**, 3772–3779 (2008)
256. L. Martiradonna, T. Stomeo, M.D. Giorgi, R. Cingolani, M.D. Vittorio, Nanopatterning of colloidal nanocrystals emitters dispersed in a PMMA matrix by e-beam lithography. *Microelectron. Eng.* **83**, 1478–1481 (2006)
257. L. Martiradonna, T. Stomeo, L. Carbone, G. Morello, A. Salhi, M. De Giorgi, R. Cingolani, M. De Vittorio, Nanopositioning of colloidal nanocrystal emitters by means of photolithography and e-beam lithography. *Phys. Status Solidi B* **243**, 3972–3975 (2006)
258. S.G. Lukishova, L.J. Bissell, J. Winkler, C.R. Stroud, Resonance in quantum dot fluorescence in a photonic bandgap liquid crystal host. *Opt. Lett.* **37**, 1259–1261 (2012)
259. R. Albrecht, A. Bommer, C. Deutsch, J. Reichel, C. Becher, Coupling of a single nitrogen-vacancy center in diamond to a fiber-based microcavity. *Phys. Rev. Lett.* **110**, 243602 (2013)
260. H. Kaupp, C. Deutsch, H.-C. Chang, J. Reichel, T.W. Hänsch, D. Hunger, Scaling laws of the cavity enhancement for nitrogen-vacancy centers in diamond. *Phys. Rev. A* **88**, 53812 (2013)

261. R. Albrecht, A. Bommer, C. Pauly, F. Mücklich, A.W. Schell, P. Engel, T. Schröder, O. Benson, J. Reichel, C. Becher, Narrow-band single photon emission at room temperature based on a single nitrogen-vacancy center coupled to an all-fiber-cavity. *Appl. Phys. Lett.* **105**, 73113 (2014)
262. S. Johnson, P.R. Dolan, T. Grange, A.A.P. Trichet, G. Hornecker, Y.C. Chen, L. Weng, G. M. Hughes, A.A.R. Watt, A. Auffèves, J.M. Smith, Tunable cavity coupling of the zero phonon line of a nitrogen-vacancy defect in diamond. *New J. Phys.* **17**, 122003 (2015)
263. A. Chiasera, Y. Dumeige, P. Féron, M. Ferrari, Y. Jestin, G. Nunzi Conti, S. Pelli, S. Soria, G.C. Righini, Spherical whispering-gallery-mode microresonators. *Laser Photonics Rev.* **4**, 457–482 (2010)
264. A.B. Matsko, V.S. Ilchenko, Optical resonators with whispering gallery modes I: basics. *IEEE J. Sel. Top. Quantum Electron.* **12**, 3–14 (2006)
265. A.B. Matsko, V.S. Ilchenko, Optical resonators with whispering gallery modes II: applications. *IEEE J. Sel. Top. Quantum Electron.* **12**, 15–32 (2006)
266. Y.P. Rakovich, J.F. Donegan, Photonic atoms and molecules. *Laser Photonics Rev.* **4**, 179–191 (2010)
267. J. Ward, O. Benson, WGM microresonators: sensing, lasing and fundamental optics with microspheres. *Laser Photonics Rev.* **5**, 553–570 (2011)
268. T.J. Kippenberg, Nonlinear optics in ultra-high-Q whispering-gallery optical microcavities. Ph.D. thesis (California Institute of Technology, Pasadena, CA, 2004)
269. M.V. Artemyev, U. Woggon, R. Wannemacher, H. Jaschinski, W. Langbein, Light trapped in a photonic dot: microspheres act as a cavity for quantum dot emission. *Nano Lett.* **1**, 309–314 (2001)
270. S. Schietinger, O. Benson, Coupling single NV-centres to high-Q whispering gallery modes of a preselected frequency-matched microresonator. *J. Phys. B At. Mol. Opt. Phys.* **42**, 114001 (2009)
271. M. Larsson, K.N. Dinyari, H. Wang, Composite optical microcavity of diamond nanopillar and silica microsphere. *Nano Lett.* **9**, 1447–1450 (2009)
272. A.W. Schell, Photonic applications and hybrid integration of single nitrogen vacancy centres in nanodiamond. Ph.D thesis (Humboldt University, Berlin, 2014)
273. M. Gregor, R. Henze, T. Schröder, O. Benson, On-demand positioning of a preselected quantum emitter on a fiber-coupled toroidal microresonator. *Appl. Phys. Lett.* **95**, 153110 (2009)
274. A.W. Schell, J. Kaschke, J. Fischer, R. Henze, J. Wolters, M. Wegener, O. Benson, Three-dimensional quantum photonic elements based on single nitrogen vacancy-centres in laser-written microstructures. *Sci. Rep.* **3**, 1577 (2013)
275. S.A. Maier, *Plasmonics: Fundamentals and Applications* (Springer Science & Business Media, 2007)
276. V. Klimov, *Nanoplasmonics* (CRC Press, 2014)
277. M. Agio, A. Alù, *Optical Antennas* (Cambridge University Press, 2013)
278. P. Bharadwaj, Antenna-coupled photoemission from single quantum emitters. Ph.D. thesis (University of Rochester, Rochester, NY, 2012)
279. P. Bharadwaj, B. Deutscher, L. Novotny, Optical antennas. *Adv. Opt. Photonics* **1**, 438 (2009)
280. L. Novotny, N. van Hulst, Antennas for light. *Nat. Photonics* **5**, 83–90 (2011)
281. A.E. Krasnok, I.S. Maksymov, A.I. Denisyuk, P.A. Belov, A.E. Miroshnichenko, C.R. Simovski, Yu.S. Kivshar, Optical nanoantennas. *Physics-Uspekhi* **56**(6), 539 (2013)
282. B. Hecht, P. Mühlischlegel, J.N. Farahani, H.-J. Eisler, D.W. Pohl, Resonant optical antennas and single emitters, in *Tip Enhancement*, ed. by S. Kawata, V.M. Shalaev (Elsevier, 2011), pp. 275–307
283. C. Ciraci, R.T. Hill, J.J. Mock, Y. Urzumov, A.I. Fernández-Domínguez, S.A. Maier, J.B. Pendry, A. Chilkoti, D.R. Smith, Probing the ultimate limits of plasmonic enhancement. *Science* **337**, 1072–1074 (2012)

284. A. Kinkhabwala, Z. Yu, S. Fan, Y. Avlasevich, K. Müllen, W.E. Moerner, Large single-molecule fluorescence enhancements produced by a bowtie nanoantenna. *Nat. Photonics* **3**, 654–657 (2009)
285. G.M. Akselrod, C. Argyropoulos, T.B. Hoang, C. Ciraci, C. Fang, J. Huang, D.R. Smith, M.H. Mikkelsen, Probing the mechanisms of large Purcell enhancement in plasmonic nanoantennas. *Nat. Photonics* **8**, 835–840 (2014)
286. T.B. Hoang, G.M. Akselrod, C. Argyropoulos, J. Huang, D.R. Smith, M.H. Mikkelsen, Ultrafast spontaneous emission source using plasmonic nanoantennas. *Nat. Commun.* **6**, Article number 7788 (2015)
287. A. Rose, T.B. Hoang, F. McGuire, J.J. Mock, C. Ciraci, D.R. Smith, M.H. Mikkelsen, Control of radiative processes using tunable plasmonic nanopatch antennas. *Nano Lett.* **14**, 4797–4802 (2014)
288. T.B. Hoang, G.M. Akselrod, M.H. Mikkelsen, Ultrafast room-temperature single photon emission from quantum dots coupled to plasmonic nanocavities. *Nano Lett.* **16**, 270–275 (2016)
289. P. Anger, P. Bharadwaj, L. Novotny, Enhancement and quenching of single-molecule fluorescence. *Phys. Rev. Lett.* **96**, 113002–113004 (2006)
290. L. Rogobete, F. Kaminski, M. Agio, V. Sandoghdar, Design of plasmonic nanoantennae for enhancing spontaneous emission. *Opt. Lett.* **32**, 1623–1625 (2007)
291. M. Agio, Optical antennas as nanoscale resonators. *Nanoscale* **4**, 692–706 (2012)
292. M. Barth, N. Nusse, B. Lochel, O. Benson, Controlled coupling of a single-diamond nanocrystal to a photonic crystal cavity. *Opt. Lett.* **34**, 1108–1110 (2009)
293. T. Junno, K. Deppert, L. Montelius, L. Samuelson, Controlled manipulation of nanoparticles with an atomic force microscope. *Appl. Phys. Lett.* **66**, 3627–3629 (1995)
294. M.S. Anderson, Nearfield surface enhanced spectroscopy using targeted nanoparticle deposition. *Appl. Phys. Lett.* **92**, 123101 (2008)
295. J. Wolters, G. Kewes, A.W. Schell, N. Nüsse, M. Schoengen, B. Löchel, T. Hanke, R. Bratschitsch, A. Leitenstorfer, T. Aichele, O. Benson, Coupling of single nitrogen-vacancy defect centers in diamond nanocrystals to optical antennas and photonic crystal cavities. *Phys. Status Solidi B* **249**, 918–924 (2012)
296. D. Dregely, K. Lindfors, J. Dorfmüller, M. Hentschel, M. Becker, J. Wrachtrup, M. Lippitz, R. Vogelgesang, H. Giessen, Plasmonic antennas, positioning, and coupling of individual quantum systems. *Phys. Status Solidi B* **249**, 666–677 (2012)
297. Y. Alaverdyan, N. Vamivakas, J. Barnes, C. Lebouteiller, J. Hare, M. Atatüre, Spectral tunability of a plasmonic antenna with a dielectric nanocrystal. *Opt. Express* **19**, 18175 (2011)
298. M.G. Harats, N. Livneh, R. Rapaport, Design, fabrication and characterization of a hybrid metal-dielectric nanoantenna with a single nanocrystal for directional single photon emission. *Opt. Mater. Express* **7**(3), 834–843 (2017)
299. N. Livneh, M.G. Harats, D. Istrati, H.S. Eisenberg, R. Rapaport, Highly directional room-temperature single photon device. *Nano Lett.* **16**, 2527–2532 (2016)
300. E. Dulkeith, T. Niedereichholz, T.A. Klar, J. Feldmann, G. von Plessen, D.I. Gittins, K.S. Mayya, F. Caruso, Plasmon emission in photoexcited gold nanoparticles. *Phys. Rev. B* **70**, 205424 (2004)
301. G.T. Boyd, Z.H. Yu, Y.R. Shen, Photoinduced luminescence from the noble metals and its enhancement on roughened surfaces. *Phys. Rev. B* **33**, 7923–7936 (1986)
302. M.R. Beversluis, A. Bouhelier, L. Novotny, Continuum generation from single gold nanostructures through near-field mediated intraband transitions. *Phys. Rev. B* **68**, 115433 (2003)
303. Z. Guo, Y. Zhang, Y. DuanMu, L. Xu, S. Xie, N. Gu, Facile synthesis of micrometer-sized gold nanoplates through an aniline-assisted route in ethylene glycol solution. *Colloids Surf. Physicochem. Eng. Asp.* **278**, 33–38 (2006)

304. J.-S. Huang, V. Callegari, P. Geisler, C. Brüning, J. Kern, J.C. Prangsma, X. Wu, T. Feichtner, J. Ziegler, P. Weinmann, M. Kamp, A. Forchel, P. Biagioni, U. Sennhauser, B. Hecht, Atomically flat single-crystalline gold nanostructures for plasmonic nanocircuitry. *Nat. Commun.* **1**, 150 (2010)
305. K.J. Russell, T.-L. Liu, S. Cui, E.L. Hu, Large spontaneous emission enhancement in plasmonic nanocavities. *Nat. Photonics* **6**, 459–462 (2012)
306. K.J. Russell, E.L. Hu, Gap-mode plasmonic nanocavity. *Appl. Phys. Lett.* **97**, 163115 (2010)
307. R. Esteban, T.V. Teperik, J.J. Greffet, Optical patch antennas for single photon emission using surface plasmon resonances. *Phys. Rev. Lett.* **104** (2010)
308. C. Belacel, B. Habert, F. Bigourdan, F. Marquier, J.-P. Hugonin, S. Michaelis de Vasconcellos, X. Lafosse, L. Coolen, C. Schwob, C. Javaux, B. Dubertret, J.-J. Greffet, P. Senellart, A. Maitre, Controlling spontaneous emission with plasmonic optical patch antennas. *Nano Lett.* **13**, 1516–1521 (2013)
309. M. Yi, D. Zhang, P. Wang, X. Jiao, S. Blair, X. Wen, Q. Fu, Y. Lu, H. Ming, Plasmonic interaction between silver nano-cubes and a silver ground plane studied by surface-enhanced Raman scattering. *Plasmonics* **6**, 515–519 (2011)
310. J.B. Lassiter, F. McGuire, J.J. Mock, C. Ciraci, R.T. Hill, B.J. Wiley, A. Chilkoti, D.R. Smith, Plasmonic waveguide modes of film-coupled metallic nanocubes. *Nano Lett.* **13**, 5866–5872 (2013)
311. C.T. Yuan, Y.C. Wang, H.W. Cheng, H.S. Wang, M.Y. Kuo, M.H. Shih, J. Tang, Modification of fluorescence properties in single colloidal quantum dots by coupling to plasmonic gap modes. *J. Phys. Chem. C* **117**, 12762–12768 (2013)
312. S.P. Eliseev, A.G. Vitukhnovsky, D.A. Chubich, N.S. Kurochkin, V.V. Sychev, A.A. Marchenko, Picosecond time of spontaneous emission in plasmonic patch nanoantennas. *JETP Lett.* **103**(2), 82–86 (2016)
313. A. Moreau, C. Ciraci, J.J. Mock, R.T. Hill, Q. Wang, B.J. Wiley, A. Chilkoti, D.R. Smith, Controlled-reflectance surfaces with film-coupled colloidal nanoantennas. *Nature* **492**, 86–89 (2012)
314. A.J. Hoffman, L. Alekseyev, S.S. Howard, K.J. Franz, D. Wasserman, V.A. Podolskiy, E.E. Narimanov, D.L. Sivco, C. Gmachl, Negative refraction in semiconductor metamaterials. *Nat. Mater.* **6**, 946–950 (2007)
315. J. Yao, Z. Liu, Y. Liu, Y. Wang, C. Sun, G. Bartal, A.M. Stacy, X. Zhang, Optical negative refraction in bulk metamaterials of nanowires. *Science* **321**, 930–930 (2008)
316. M.A. Noginov, Y.A. Barnakov, G. Zhu, T. Tumkur, H. Li, E.E. Narimanov, Bulk photonic metamaterial with hyperbolic dispersion. *Appl. Phys. Lett.* **94**, 151105 (2009)
317. M.A. Noginov, H. Li, Y.A. Barnakov, D. Dryden, G. Nataraj, G. Zhu, C.E. Bonner, M. Mayy, Z. Jacob, E.E. Narimanov, Controlling spontaneous emission with metamaterials. *Opt. Lett.* **35**, 1863–1865 (2010)
318. A. Poddubny, I. Iorsh, P. Belov, Y. Kivshar, Hyperbolic metamaterials. *Nat. Photonics* **7**, 948–957 (2013)
319. M.Y. Shalaginov, S. Bogdanov, V.V. Vorobyov, A.S. Lagutchev, A.V. Akimov, A. Boltasseva, V.M. Shalaev, Enhancement of single-photon sources with metamaterials, in *From Atomic to Mesoscale: The Role of Quantum Coherence in Systems of Various Complexities*, ed. by S.A. Malinovskaya, I. Novikova (World Scientific, 2015), pp. 123–148
320. Z. Jacob, I.I. Smolyaninov, E.E. Narimanov, Broadband Purcell effect: radiative decay engineering with metamaterials. *Appl. Phys. Lett.* **100**, 181105 (2012)
321. C.L. Cortes, W. Newman, S. Molesky, Z. Jacob, Quantum nanophotonics using hyperbolic metamaterials. *J. Opt.* **14**, 63001 (2012)
322. J. Kim, V.P. Drachev, Z. Jacob, G.V. Naik, A. Boltasseva, E.E. Narimanov, V.M. Shalaev, Improving the radiative decay rate for dye molecules with hyperbolic metamaterials. *Opt. Express* **20**, 8100–8116 (2012)

323. Z. Jacob, J.-Y. Kim, G.V. Naik, A. Boltasseva, E.E. Narimanov, V.M. Shalaev, Engineering photonic density of states using metamaterials. *Appl. Phys. B* **100**, 215–218 (2010)
324. M.Y. Shalaginov, V.V. Vorobyov, J. Liu, M. Ferrera, A.V. Akimov, A. Lagutchev, A.N. Smolyaninov, V.V. Klimov, J. Irudayaraj, A.V. Kildishev, A. Boltasseva, V.M. Shalaev, Enhancement of single-photon emission from nitrogen-vacancy centers with TiN/(Al, Sc)N hyperbolic metamaterial. *Laser Photonics Rev.* **9**, 120–127 (2015)
325. G.V. Naik, J.L. Schroeder, X. Ni, A.V. Kildishev, T.D. Sands, A. Boltasseva, Titanium nitride as a plasmonic material for visible and near-infrared wavelengths. *Opt. Mater. Express* **2**, 478–489 (2012)
326. G.V. Naik, B. Saha, J. Liu, S.M. Saber, E.A. Stach, J.M.K. Irudayaraj, T.D. Sands, V.M. Shalaev, A. Boltasseva, Epitaxial superlattices with titanium nitride as a plasmonic component for optical hyperbolic metamaterials. *Proc. Natl. Acad. Sci.* **111**, 7546–7551 (2014)
327. M.Y. Shalaginov, S. Ishii, J. Liu, J. Liu, J. Irudayaraj, A. Lagutchev, A.V. Kildishev, V.M. Shalaev, Broadband enhancement of spontaneous emission from nitrogen-vacancy centers in nanodiamonds by hyperbolic metamaterials. *Appl. Phys. Lett.* **102**, 173114 (2013)
328. H.N.S. Krishnamoorthy, Z. Jacob, E. Narimanov, I. Kretzschmar, V.M. Menon, Topological transitions in metamaterials. *Science* **336**, 205–209 (2012)
329. K.V. Sreekanth, K.H. Krishna, A.D. Luca, G. Strangi, Large spontaneous emission rate enhancement in grating coupled hyperbolic metamaterials. *Sci. Rep.* **4**, 6340 (2014)
330. D. Lu, J.J. Kan, E.E. Fullerton, Z. Liu, Enhancing spontaneous emission rates of molecules using nanopatterned multilayer hyperbolic metamaterials. *Nat. Nanotechnol.* **9**, 48–53 (2014)
331. T. Galfsky, H.N.S. Krishnamoorthy, W. Newman, E.E. Narimanov, Z. Jacob, V.M. Menon, Active hyperbolic metamaterials: enhanced spontaneous emission and light extraction. *Optica* **2**, 62–65 (2015)
332. T. Galfsky, J. Gu, E.E. Narimanov, V.M. Menon, Photonic hypercrystals for control of light -matter interactions. *PNAS* **114**, 5125–5129 (2017)
333. S.G. Lukishova, A.C. Liapis, L.J. Bissell, G.M. Gehring, R.W. Boyd, Single-photon experiments with liquid crystals for quantum science and quantum engineering applications. *Liq. Cryst. Rev.* **2**, 111–129 (2014)
334. S.G. Lukishova, R.W. Boyd, C.R. Stroud, Efficient room-temperature source of polarized single photons. <http://www.google.com/patents/US7253871> (2007)
335. S.G. Lukishova, A.W. Schmid, C.M. Supranowitz, N. Lippa, A.J. McNamara, R.W. Boyd, C.R. Stroud, Dye-doped cholesteric-liquid-crystal room-temperature single-photon source. *J. Mod. Opt.* **51**, 1535–1547 (2004)
336. S. Lukishova, R. Knox, P. Freivald, A. McNamara, R. Boyd, C. Stroud, A. Schmid, K. Marshall, Single-photon source for quantum information based on single dye molecule fluorescence in liquid crystal host. *Mol. Cryst. Liq. Cryst.* **454**, 1/[403]–14/[416] (2006)
337. S.G. Lukishova, Bissell, J. Luke, C.R. Stroud, R.W. Boyd, Room temperature single photon sources with definite circular and linear polarizations. *Opt. Spectrosc.* **108**, 417–424 (2010)
338. S.G. Lukishova, Liquid crystals under two extremes: (1) high-power laser irradiation, and (2) single-photon level. *Mol. Cryst. Liq. Cryst.* **559**, 127–157 (2012)
339. S.G. Lukishova, A.W. Schmid, R. Knox, P. Freivald, L.J. Bissell, R.W. Boyd, C.R. Stroud, K.L. Marshall, Room temperature source of single photons of definite polarization. *J. Mod. Opt.* **54**, 417–429 (2007)
340. S.G. Lukishova, J.M. Winkler, L.J. Bissell, Quantum dot fluorescence in photonic bandgap glassy cholesteric liquid crystal structures: microcavity resonance under cw-excitation, antibunching and decay time. *Mol. Cryst. Liq. Cryst.* **595**, 98–105 (2014)
341. L. Pelliser, M. Manceau, C. Lethiec, D. Coursault, S. Vezzoli, G. Leménager, L. Coolen, M. DeVittorio, F. Pisanello, L. Carbone, A. Maitre, A. Bramati, E. Lacaze, Alignment of rod-shaped single-photon emitters driven by line defects in liquid crystals. *Adv. Funct. Mater.* **25**, 1719–1726 (2015)

342. M. Manceau: Single CdSe/CdS dot-in-rods fluorescence properties. Ph.D. thesis, University of Pierre et Marie Curie, Paris (2014). <https://hal.archives-ouvertes.fr/tel-01101939>. Accessed 25 June 2016
343. I.P. Il'chishin, E.A. Tikhonov, V.G. Tishchenko, M.T. Shpak, Generation of a tunable radiation by impurity cholesteric liquid crystals. *JETP Lett.* **32**, 24–27 (1980)
344. V.I. Kopp, B. Fan, H.K.M. Vithana, A.Z. Genack, Low-threshold lasing at the edge of a photonic stop band in cholesteric liquid crystals. *Opt. Lett.* **23**, 1707–1709 (1998)
345. L.M. Blinov, V. Chigrinov, *Electrooptic Effects in Liquid Crystal Materials* (Springer, 1996)
346. V.G. Chigrinov, V.M. Kozenkov, H.-S. Kwok, *Photoalignment of Liquid Crystalline Materials: Physics and Applications* (Wiley, 2008)
347. V.G. Chigrinov, Photoaligning and photopatterning—a new challenge in liquid crystal photonics. *Crystals.* **3**, 149–162 (2013)
348. P. Pieranski, Classroom experiments with chiral liquid crystals, in *Chirality in Liquid Crystals*, ed. by H.-S. Kitzerow, C. Bahr (Springer, 2001), pp. 28–66
349. P. Palfy-Muhoray, W. Cao, M. Moreira, B. Taheri, A. Munoz, Photonics and lasing in liquid crystal materials. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **364**, 2747–2761 (2006)
350. H. Coles, S. Morris, Liquid-crystal lasers. *Nat. Photonics* **4**, 676–685 (2010)
351. L.M. Blinov, R. Bartolino (eds.), *Liquid Crystal Microlasers* (Transworld Research Network, Trivandrum, 2010)
352. G. Chilaya, A. Chanishvili, G. Petriashvili, R. Barberi, M.P.D. Santo, M.A. Matranga, Different approaches of employing cholesteric liquid crystals in dye lasers. *Mater. Sci. Appl.* **2**, 116 (2011)
353. S.K.H. Wei, S.H. Chen, K. Dolgaleva, S.G. Lukishova, R.W. Boyd, Robust organic lasers comprising glassy-cholesteric pentafluorene doped with a red-emitting oligofluorene. *Appl. Phys. Lett.* **94**, 41111–41113 (2009)
354. K. Dolgaleva, S.K.H. Wei, S.G. Lukishova, S.H. Chen, K. Schwertz, R.W. Boyd, Enhanced laser performance of cholesteric liquid crystals doped with oligofluorene dye. *J. Opt. Soc. Am. B.* **25**, 1496–1504 (2008)
355. T.J. Bunning, F.-H. Kreuzer, Cyclosiloxane-based liquid crystalline materials. *Trends Polym Sci.* **3**, 318–323 (1995)
356. S.G. Lukishova, A.W. Schmid, Near-field optical microscopy of defects in cholesteric oligomeric liquid crystal films. *Mol. Cryst. Liq. Cryst.* **454**, 15–21 (2006)
357. J.Y.P. Butter, B.R. Crenshaw, C. Weder, B. Hecht, Single-molecule spectroscopy of uniaxially oriented terrylene in polyethylene. *ChemPhysChem* **7**, 261–265 (2006)
358. D. Coursault, B. Zappone, A. Coati, A. Boulaoued, L. Pelliser, D. Limagne, N. Boudet, B.H. Ibrahim, A. de Martino, M. Alba, M. Goldmann, Y. Garreau, B. Gallas, E. Lacaze, Self-organized arrays of dislocations in thin smectic liquid crystal films. *Soft Matter* **12**, 678–688 (2016)
359. S.G. Lukishova, A.C. Liapis, H. Zhu, E. Hebert, K. Kuyk, S. Choudhary, R.W. Boyd, Z. Wang, L.J. Bissell, Plasmonic nanoantennas with liquid crystals for nanocrystal fluorescence enhancement and polarization selectivity of classical and quantum light sources. *Mol. Cryst. Liq. Cryst.* **657**, 173–183 (2017)
360. L. Jiang, H. Mundoor, Q. Liu, I.I. Smalyukh, Electric switching of fluorescence decay in gold-silica-dye nematic nanocolloids mediated by surface plasmons. *ACS Nano* **10**, 7064–7072 (2016)
361. R.N. Patel, T. Schröder, N. Wan, L. Li, S.L. Mouradian, E.H. Chen, D.R. Englund, Efficient photon coupling from a diamond nitrogen vacancy center by integration with silica fiber. *Light: Sci. Appl.* **5**, e16032 (2016)
362. E. Ampem-Lassen, D.A. Simpson, B.C. Gibson, S. Trpkovski, F.M. Hossain, S.T. Huntington, K. Ganeshan, L.C.L. Hollenberg, S. Prawer, Nano-manipulation of diamond-based single photon sources. *Opt. Express* **17**, 11287–11293 (2009)
363. C. Santori, D. Fattal, J. VuCkovic, G.S. Solomon, Y. Yamamoto, Indistinguishable photons from a single-photon device. *Nature* **419**, 594–597 (2002)

364. T. Legero, T. Wilk, M. Hennrich, G. Rempe, A. Kuhn, Quantum beat of two single photons. *Phys. Rev. Lett.* **93**, 70503 (2004)
365. A. Kiraz, M. Ehrl, Th. Hellerer, Ö.E. Müstecaplıoğlu, C. Bräuchle, A. Zumbusch, Indistinguishable photons from a single molecule. *Phys. Rev. Lett.* **94**, 223602 (2005)
366. R. Lettow, Y.L.A. Rezus, A. Renn, G. Zumofen, E. Ikonen, S. Götzinger, V. Sandoghdar, Quantum interference of tunably indistinguishable photons from remote organic molecules. *Phys. Rev. Lett.* **104**, 123605 (2010)
367. E.B. Flagg, A. Muller, S.V. Polyakov, A. Ling, A. Migdall, G.S. Solomon, Interference of single photons from two separate semiconductor quantum dots. *Phys. Rev. Lett.* **104**, 137401 (2010)
368. R.B. Patel, A.J. Bennett, I. Farrer, C.A. Nicoll, D.A. Ritchie, A.J. Shields, Two-photon interference of the emission from electrically tunable remote quantum dots. *Nat. Photonics* **4**, 632–635 (2010)
369. A. Sipahigil, M.L. Goldman, E. Togan, Y. Chu, M. Markham, D.J. Twitchen, A.S. Zibrov, A. Kubanek, M.D. Lukin, Quantum interference of single photons from remote nitrogen-vacancy centers in diamond. *Phys. Rev. Lett.* **108**, 143601 (2012)
370. A.V. Kuhlmann, J.H. Prechtel, J. Houel, A. Ludwig, D. Reuter, A.D. Wieck, R.J. Warburton, Transform-limited single photons from a single quantum dot. *Nat. Commun.* **6**, 8204 (2015)
371. C. Matthiesen, A.N. Vamivakas, M. Atatüre, Subnatural linewidth single photons from a quantum dot. *Phys. Rev. Lett.* **108**, 93602 (2012)
372. C. Matthiesen, M. Geller, C.H.H. Schulte, C. Le Gall, J. Hansom, Z. Li, M. Hugues, E. Clarke, M. Atatüre, Phase-locked indistinguishable photons with synthesized waveforms from a solid-state source. *Nat. Commun.* **4**, 1600 (2013)
373. O. Gazzano, S. Michaelis de Vasconcellos, C. Arnold, A. Nowak, E. Galopin, I. Sagnes, L. Lanco, A. Lemaître, P. Senellart, Bright solid-state sources of indistinguishable single photons. *Nat. Commun.* **4**, 1425 (2013)
374. X. Ding, Y. He, Z.-C. Duan, N. Gregersen, M.-C. Chen, S. Unsleber, S. Maier, C. Schneider, M. Kamp, S. Höfling, C.-Y. Lu, J.-W. Pan, On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar. *Phys. Rev. Lett.* **116**, 20401 (2016)
375. H. Wang, Z.-C. Duan, Y.-H. Li, S. Chen, J.-P. Li, Y.-M. He, M.-C. Chen, Y. He, X. Ding, C.-Z. Peng, C. Schneider, M. Kamp, S. Höfling, C.-Y. Lu, J.-W. Pan, Near-transform-limited single photons from an efficient solid-state quantum emitter. *Phys. Rev. Lett.* **116**, 213601 (2016)
376. W.T. Buttler, R.J. Hughes, P.G. Kwiat, S.K. Lamoreaux, G.G. Luther, G.L. Morgan, J.E. Nordholz, C.G. Peterson, C.M. Simmons, Practical free-space quantum key distribution over 1 km. *Phys. Rev. Lett.* **81**, 3283–3286 (1998)
377. X. Shan, X. Sun, J. Luo, Z. Tan, M. Zhan, Free-space quantum key distribution with Rb vapor filters. *Appl. Phys. Lett.* **89**, 191121 (2006)
378. D.N. Wolf, Ultra-bright single photon source. PSI—Physical Sciences Inc. <http://www.psicorp.com/content/ultra-bright-single-photon-source>
379. N. Sangouard, H. Zbinden, What are single photons good for? *J. Mod. Opt.* **59**, 1458–1464 (2012)
380. A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, V. Scarani, Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.* **98**, 230501 (2007)
381. N. Gisin, S. Pironio, N. Sangouard, Proposal for implementing device-independent quantum key distribution based on a heralded qubit amplifier. *Phys. Rev. Lett.* **105**, 70501 (2010)
382. J. Minář, H. de Riedmatten, N. Sangouard, Quantum repeaters based on heralded qubit amplifiers. *Phys. Rev. A* **85**, 32313 (2012)
383. C. Simon, H. de Riedmatten, M. Afzelius, N. Sangouard, H. Zbinden, N. Gisin, Quantum repeaters with photon pair sources and multimode memories. *Phys. Rev. Lett.* **98**, 190503 (2007)

384. T. Grange, G. Hornecker, D. Hunger, J.-P. Poizat, J.-M. Gérard, P. Senellart, A. Auffèves, Cavity-funneled generation of indistinguishable single photons from strongly dissipative quantum emitters. *Phys. Rev. Lett.* **114**, 193601 (2015)
385. S. Wein, N. Lauk, R. Ghobadi, C. Simon, Feasibility of efficient room-temperature solid-state sources of indistinguishable single photons using ultrasmall mode volume cavities. *Phys. Rev. B* **97**, 205418 (2018)
386. B. Gurlek, V. Sandoghdar, D. Martín-Cano, Manipulation of quenching in nanoantenna-emitter systems enabled by external detuned cavities: a path to enhance strong-coupling. *ACS Photonics* **5**, 456–461 (2018)
387. H. Siampour, S. Kumar, S.I. Bozhevolnyi, Chip-integrated plasmonic cavity-enhanced single nitrogen-vacancy center emission. *Nanoscale* **9**, 17902–17908 (2017)
388. A. Singh, J.T. Hugall, G. Calbris, N.F. van Hulst, Fiber-based optical nanoantennas for single-molecule imaging and sensing. *J. Light. Technol.* **33**, 2371–2377 (2015)
389. X. Wang, G. Venugopal, J. Zeng, Y. Chen, D.H. Lee, N.M. Litchinitser, A.N. Cartwright, Optical fiber metamagnetics. *Opt. Express* **19**, 19813–19821 (2011)
390. M.A. Noginov, L. Gu, J. Livenere, G. Zhu, A.K. Pradhan, R. Mundle, M. Bahoura, Y.A. Barnakov, V.A. Podolskiy, Transparent conductive oxides: plasmonic materials for telecom wavelengths. *Appl. Phys. Lett.* **99**, 21101 (2011)
391. S. Zaske, A. Lenhard, C.A. Keßler, J. Kettler, C. Hepp, C. Arend, R. Albrecht, W.-M. Schulz, M. Jetter, P. Michler, C. Becher, Visible-to-Telecom Quantum frequency conversion of light from a single quantum emitter. *Phys. Rev. Lett.* **109**, 147404 (2012)
392. M.T. Rakher, L. Ma, O. Slattery, X. Tang, K. Srinivasan, Quantum transduction of telecommunications-band single photons from a quantum dot by frequency upconversion. *Nat. Photonics* **4**, 786–791 (2010)
393. S. Ates, I. Agha, A. Gulinatti, I. Rech, M.T. Rakher, A. Badolato, K. Srinivasan, Two-photon interference using background-free quantum frequency conversion of single photons emitted by an InAs quantum dot. *Phys. Rev. Lett.* **109**, 147405 (2012)
394. J. Demas, P. Steinvurzel, B. Tai, L. Rishøj, Y. Chen, S. Ramachandran, Intermodal nonlinear mixing with Bessel beams in optical fiber. *Optica* **2**, 14–17 (2015)
395. G. Gibson, J. Courtial, M.J. Padgett, M. Vasnetsov, V. Pas'ko, S.M. Barnett, S. Franke-Arnold, Free-space information transfer using light beams carrying orbital angular momentum. *Opt. Express* **12**, 5448–5456 (2004)
396. J. Wang, J.-Y. Yang, I.M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, A.E. Willner, Terabit free-space data transmission employing orbital angular momentum multiplexing. *Nat. Photonics* **6**, 488–496 (2012)
397. S. Gröblacher, T. Jennewein, A. Vaziri, G. Weihs, A. Zeilinger, Experimental quantum cryptography with qutrits. *New J. Phys.* **8**, 75 (2006)
398. M. Mirhosseini, O.S. Magaña-Loaiza, M.N. O'Sullivan, B. Rodenburg, M. Malik, M.P.J. Lavery, M.J. Padgett, D.J. Gauthier, R.W. Boyd, High-dimensional quantum cryptography with twisted light. *New J. Phys.* **17**, 33033 (2015)
399. H. Choi, D. Zhu, Y. Yoon, D. Englund, Indistinguishable single-photon sources with dissipative emitter coupled to cascaded cavities. arXiv: 1809.01645, September 2018
400. T. Iwasaki, Y. Miyamoto, T. Taniguchi, P. Siyushev, M.H. Metsch, F. Jelezko, M. Hatano, Tin-vacancy quantum emitters in diamond, *Phys. Rev. Lett.* **119**, 253601 (2017)
401. S.I. Bogdanov, M.Y. Shalaginov, A.S. Lagutchev, C.-C. Chiang, D. Shah, A.S. Baburin, I.A. Ryzhikov, I.A. Rodionov, A.V. Kildishev, A. Boltasseva, V.M. Shalaev, Ultrabright room-temperature sub-nanosecond emission from single nitrogen-vacancy centers coupled to nanopatch antennas, *Nano Lett.* **18**, 4837–4844 (2018)
402. S. Bogdanov, M.Y. Shalaginov, A. Boltasseva, V.M. Shalaev, Material platforms for integrated quantum photonics. *Opt. Mater. Express* **7**, 111–132 (2017)

Chapter 5

Time-Multiplexed Methods for Optical Quantum Information Processing



**Michelle Victora, Fumihiro Kaneda, Fedor Bergmann, Jia Jun Wong,
Austin Graf and Paul Kwiat**

Abstract Quantum information processing with photons can be greatly enhanced by incorporating time-multiplexing methods. Not only can time-bin encoding be very useful in its own right, multiplexing techniques can lead to more efficient single- and multi-photon sources, improved detectors, and high-bandwidth quantum memories, as well as enhanced applications such as quantum random walks and entanglement swapping. Here we present an overview of some of the methods used and the results achievable when explicitly using the time degree of freedom of photons.

M. Victora (✉) · P. Kwiat (✉)

Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL
61801-3080, USA

e-mail: victora2@illinois.edu

P. Kwiat

e-mail: kwiat@illinois.edu

F. Kaneda

Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai,
Miyagi 980-8578, Japan

e-mail: fumikaneda@gmail.com

F. Bergmann

Bergmann Messgeräte Entwicklung KG, Murnau
82418, Germany

e-mail: fbergmann@bme-bergmann.de

J. J. Wong

DSO National Laboratories, 12 Science Park Drive, Singapore
S118225, Singapore

e-mail: WJiaJun@dso.org.sg

A. Graf

The Institute of Optics, University of Rochester, Rochester, NY
14627, USA

e-mail: agraf@ur.rochester.edu

5.1 Introduction

Optical quantum information processing is now a relatively mature field, having seen demonstrations of quantum key distribution, teleportation, dense coding, entanglement swapping, quantum-enhanced metrology, and basic quantum computing. In many cases, the quantum information has been stored in the polarization states or optical path (spatial mode) of the photons. However, it is also possible to use the time degree of freedom for information processing [1]. This has some advantages in that it can be more robust to atmospheric disturbance or propagation through fibers, whose properties are effectively fixed over the nanosecond timescales of the encoding. Also, one in general does not need a large array of detectors (as one would for multiple spatial-mode states); instead, it suffices that the single-photon detectors are able to operate continuously, i.e., a single detector can easily detect many distinguishable time bins. Moreover, as we will see below, there can be a large enhancement in resources by incorporating time multiplexing. Single-photon sources can be made more efficient, with a lower probability of unwanted multi-photon production. Conversely, the ability to produce higher-order photon states (e.g., Fock states) can be greatly enhanced compared to non-multiplexing techniques. Similarly, detectors can also be improved using such multiplexing techniques—non-photon number resolving detectors can nevertheless display effective photon number resolution by using time-multiplexing techniques to split an incident pulse into several resolvable time bins. There are also a number of applications that have been explored using time-multiplexing techniques, including quantum random walks, entanglement swapping, and even computation using a single spatial mode. Although it is not unlikely that any future optical quantum information systems will use a variety of degrees of freedom simultaneously (e.g., favoring hyper-entanglement over simple single-degree of freedom entanglement [2]), these techniques are nevertheless valuable in their own right. Here we give an overview of some of the achievements in this area.

5.2 Time-Multiplexed Heralded Single-Photon Source

5.2.1 *Introduction—Heralded Single-Photon Source*

One of the key technologies to scale up quantum information processing is efficient (and ideally deterministic) preparation of quantum states of light such as single-photon and definite multi-photon states. There are essentially two approaches to generate single-photon states: single-emitter and nonlinear-optics approaches [3]. Single emitters such as trapped single atoms [4, 5] and ions [6] and semiconductor sources (e.g., quantum dots [7, 8] and color centers in diamonds [9, 10]) do generate true single-photon states, and the semiconductor systems in particular have great potential for integration. However, most single emitters generate single photons in a narrow bandwidth (5–100 MHz) that may be unsuitable for high-speed applica-

tions and integrated optics. For example, the maximum generation rate of a 5-MHz bandwidth source is ~ 10 MHz, assuming non-overlapping single-photon Gaussian wave packets. Moreover, for semiconductor sources, generating “indistinguishable” photons remains a challenge¹: photons from different sources should be in pure states with identical emission spectra and spatial modes so that they can exhibit the multi-photon interference [15] required by many quantum information protocols.

The other approach uses photon-pair generation via nonlinear optical processes such as spontaneous parametric down-conversion (SPDC) and spontaneous four-wave mixing (SFWM), which have been widely used for generating entangled photon pairs and multi-photon quantum states and for implementing small-scale quantum information protocols [16]. A photon pair generated by nonlinear optics can be used to produce a “heralded” single-photon state: detecting one of the photons heralds the other’s presence in a well-defined spatio-temporal mode [17]. Moreover, the current technology is reaching almost perfect collection efficiency of photon-pair states [18–20] and single-photon indistinguishability [21–23]. However, these nonlinear optical processes are probabilistic: one cannot obtain a photon pair or heralded single photon on demand. Unfortunately, one cannot indefinitely increase the generation efficiency and produce mean number of photon pairs p simply by increasing the pumping power, because that also increases the likelihood of unwanted k -pair generation as p^k . For example, state-of-the-art 6-photon experiments have typical detected rates of only ~ 3 per second [24], and recent 8-photon experiments [25, 26] had a final detected rate below 10 per hour! This approach is clearly not scalable.

In this section, we describe time multiplexing of a heralded single-photon source, a way to efficiently overcome the probabilistic nature of the heralded single-photon generation. By using a periodically pumped heralded single-photon source and an optical delay line that is switched according to a heralding signal of photon generation, the probability of generating a single-photon state can be multiplexed and concentrated into a single output time window. We briefly describe the working principle of such a time-multiplexed source, followed by recent experimental implementations [27] that achieved substantial enhancement in the single-photon generation probability. Finally, we discuss possible improvements of the source toward near deterministic generation.

5.2.2 Working Principle

Time-multiplexing² of a heralded single-photon source was first proposed by Pittman, Jacobs, and Franson [33] in 2002. The method has since been extended

¹Indistinguishability of sequentially generated photons from a single semiconductor quantum dot has been reported [11–14].

²A similar technique based on “spatial” multiplexing has also been proposed [28–32], but this is much more resource intensive: a spatial-multiplexed source analogous to the time-multiplexed source described here would need ~ 30 photon pair sources (either using independent crystals or by extracting multiple photon-pair sources from a single crystal), low-loss binary (2 to 1) switchyard elements, and detectors.

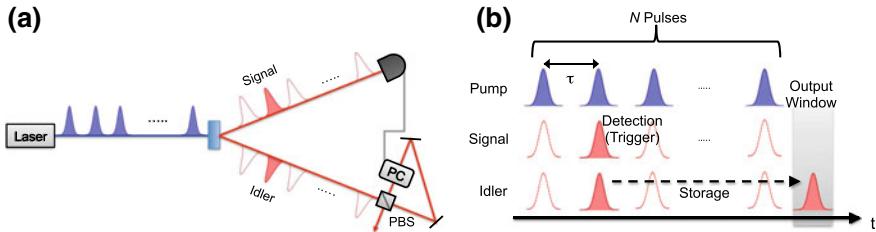


Fig. 5.1 **a** Simplified schematic diagram of a time-multiplexed heralded single-photon source. **b** Timing diagram of pump, signal, and idler photons in the time multiplexing. *SPD* single-photon detector, *PC* Pockels cell, *PBS* polarizing beam splitter. From [40]

and analyzed in great detail [34–39]. Figure 5.1 shows the basic idea; a laser pulse train containing N pulses with a period τ pumps a nonlinear crystal, and generates photon pairs (i.e., signal and idler photons). Using a sufficiently large number N , one can obtain photon pairs in one or more time slots. Signal photons generated are sent to a single-photon detector (SPD) whose firing heralds in which time slot the corresponding idler photon is present. By using an adjustable storage cavity with a matched cavity length τ , any of the time slots heralded to contain an idler photon can be multiplexed onto a single output time window: when the idler photon enters the cavity via a polarizing beam-splitter (PBS), a Pockels cell (PC) in the cavity is fired by a trigger signal from the SPD, rotating the photon polarization by 90°, thus storing the photon in the cavity. The idler photon is then released from the cavity in the desired output time slot by a second switching of the PC. Consequently, one can increase the multiplexed single-photon probability $P_M(1)$ during the output time window according to the number of pump pulses (time slots) N used for one cycle of the multiplexing; in a lossless system this probability can be made arbitrarily close to 1, at the cost of a lower final photon repetition rate. Moreover, if N is large, the probability of generating unwanted multiple pairs in a given time slot can be made arbitrarily small, because the pump energy is “diluted” over the N time slots. Hence, for a lossless SPDC source and storage cavity, this time-multiplexing technique works as a pseudo-deterministic single-photon source.

Theoretical details of such a time-multiplexed source, accounting for losses and inefficiencies of components (SPD, PC, SPDC source and other optics) are discussed in [34, 40, 41]. As shown in [40], large (i.e., >5) enhancements to $P_M(1)$ are only possible for $\lesssim 20\%$ loss per cycle of the storage cavity. This requirement is quite challenging except for bulk-optics implementations, although many proof-of-principle demonstrations of quantum storage by light-matter interfaces [42] and integrated optics [38] have been reported.

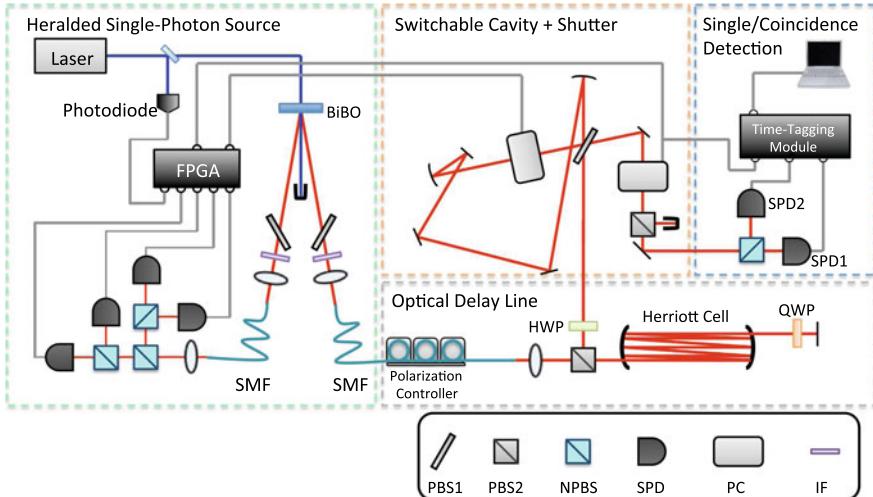


Fig. 5.2 Schematic diagram of experimental setup. BiBO: bismuth barium borate crystal. PBS1 Brewster-angled polarizing beam splitter, PBS2 cube polarizing beam splitter, NPBS non-polarizing beam splitter, SPD single-photon detector, PC Pockels cell, IF interference filter ($\Delta\lambda = 20$ nm), HWP half-wave plate, QWP quarter-wave plate, FPGA field-programmable gate array. From [40]

5.2.3 Experimental Realization

We here present one recent experimental demonstration [40] of a time-multiplexed heralded single-photon source. A schematic diagram of the experimental setup is shown in Fig. 5.2. The pump for an SPDC source was a frequency-tripled (to 355 nm) mode-locked Nd:YAG laser operating at 120 MHz. Downconversion in a bismuth barium borate (BiBO) crystal produced degenerate signal and idler pairs at 710 nm, with very high heralding efficiency into single-mode fibers (SMF) (up to 75–81% [18]). The signal photons were detected using a small “tree” of ~70%-efficient silicon avalanche photodiodes (Si-APDs), to approximate some photon-number resolution; for the case of multiple detector clicks in a time slot, the idler photons are not stored because the time slot most likely contains more than one photon. A field-programmable gate array (FPGA) module processes input signals from the SPDs, and triggers PCs at heralded and output time slots. For a switchable storage cavity, a custom Brewster-angle polarizing beamsplitter (PBS1) and PC are used, achieving high-speed (<5 ns) very low-loss (<3%) switching.

In addition to the heralded single-photon source and storage cavity, there are some key components for practical implementation. Even if the switchable cavity has very low loss, many repeated passes through it can significantly attenuate the stored idler photons. For the case that photons are heralded in multiple time slots within N cycles, photons stored later—closer to the output time window—experience less loss. One way to select only the last-produced (and therefore lowest loss) photon is to re-fire

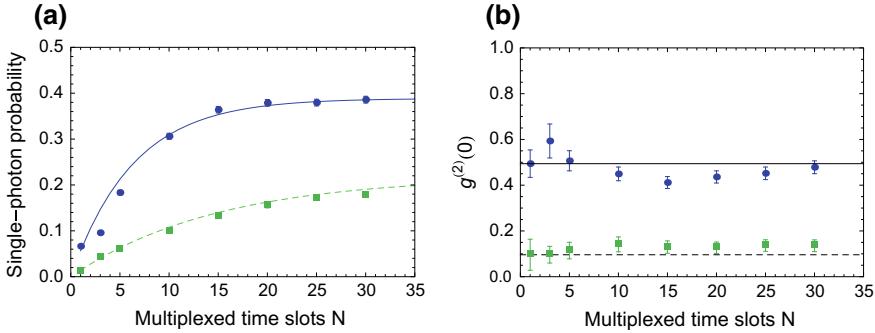


Fig. 5.3 Experimental results. **a** Single-photon probability $P_M(1)$, and **b** second-order correlation function $g^{(2)}(t = 0)$ versus number of multiplexed time slots N . Blue circles: $p = 0.35$. Green squares: $p = 0.07$. Error bars are estimated by Poissonian photon counting statistics. For (a), solid and dashed curves show theoretical predictions for the measurement with $p = 0.35$ and 0.07 , taking into account the measured experimental parameters (see [40]). Solid and dashed lines in (b) indicate the values of $g^{(2)}(0)$ measured without multiplexing ($N = 1$) for $p = 0.35$ and 0.07

the PC for each new signal photon detected; this discards any previously stored idler photon and replaces it with a “fresh” one. However, PC driver constraints severely limit this approach. Instead, a long low-loss delay line between the SPDC crystal and switchable storage cavity is introduced; this delay is long enough to contain idler photons from all N of the pump pulses. Thus, one can simply determine the time slot of the last-detected signal photon, and only switch the corresponding idler photon into the storage cavity. In [40] the delay line is implemented by a Herriot cell [43] with an optical delay of 400 ns (and transmission 84.6%), sufficiently longer than the electronic latency (120 ns) of the PC driver to hold idler photons from up to $N = 30$ cycles. A PC and PBS2 after the storage cavity are used as an “optical shutter”; this shutter opens in an output time window only if at least one heralding signal is detected within N multiplexed time slots, thereby blocking unheralded photons due to imperfect signal photon detectors and heralding efficiency. This time-multiplexed source runs with a repetition rate $R = 50$ kHz, limited by the duty cycle of the PC high-voltage driver; a more recent implementation runs upto 500 kHz [27].

Figure 5.3a shows the multiplexed single-photon probability $P_M(1)$ for various measured time slots N , estimated by

$$P_M(1) = \frac{S_1 + S_2}{R\eta_{Di}}, \quad (5.1)$$

where S_1 and S_2 are the detector count rates by SPD1 and SPD2 at the output time windows, and η_{Di} is the system-detection efficiency for the multiplexed photons. For high-power pumping (mean photon number per time slot $p = 0.35$, see blue dots in Fig. 5.3a) the multiplexed single-photon probability is increased up to $P_M(1) = 38.6 \pm 0.4\%$ by multiplexing $N = 30$, corresponding to $\sim 6\times$ enhancement compared to the non-multiplexed source ($P_M(1) = 6.8 \pm 0.4\%$ for $N = 1$). Pumping

Table 5.1 Comparison between multiplexed sources (MUX) and ideal non-multiplexed sources (NMUX). The latter assume photon-pair sources with perfect heralding efficiency and 100%-efficient detectors with photon-number-resolving functionality. For the improved multiplexed source, the values are predicted by feasible improved parameters (see Sect. 5.2.4)

		$P_M(1)$ (%)	$g^{(2)}(0)$
MUX	Previous result [40]	38.6 ± 0.4	0.479 ± 0.028
	Improved source	>80	<0.05
NMUX	Mixed state	36.8	0
	Pure state	25.0	0

with a lower power ($p = 0.07$, see green squares in Fig. 5.3a) yields an even larger ($\sim 16\times$) enhancement factor, because for $p = 0.07$, the heralding signal rate H increases almost linearly with N up to 30; in contrast, for $p = 0.35$, H is nearly saturated (i.e., close to 50 kHz) with $N \gtrsim 20$. To our knowledge, these single-photon probabilities and enhancement factors are significantly higher than all previous demonstrations of heralded single photons. Moreover, the results for $P_M(1)$ even exceed what one could achieve from an *ideal* heralded single-photon source: even using a completely lossless heralded single-photon source and 100%-efficient trigger detectors, the non-multiplexed single-photon probability is limited to 36.8% for pure states and 25.0% for mixed states³ (see Table. 5.1).

The second-order correlation function at zero time delay between the signals from SPD1 and SPD2 [45] is used to investigate the contribution of the multi-photon emissions:

$$g^{(2)}(t=0) = \frac{C_{12}H}{S_1 S_2}, \quad (5.2)$$

where C_{12} is the coincidence count rate between SPD1 and SPD2. $g^{(2)}(t=0) = (m-1)/m$ for m -photon Fock states, and sources having Poissonian photon-number statistics (e.g., coherent states) show $g^{(2)}(t=0) = 1$. For the multiplexed source, sub-Poissonian values of $g^{(2)}(t=0) < 1$ —a signature of quantum states of light—have been observed. Moreover, while the $g^{(2)}(t=0)$ value depends on p (due to the change in the likelihood of multi-photon emissions), it approximately does not depend on N : the contribution of the two-photon probability relative to the single-photon one is nearly independent of the time multiplexing (see Fig. 5.3b). These nearly constant values of $g^{(2)}(t=0)$ together with greatly enhanced single-photon probability demonstrate the capacity of time-multiplexing techniques to enhance the performance of heralded single-photon sources.

³The difference comes from the photon-number statistics of SPDC sources. For a source generating heralded photons in mixed states, the photon number distribution follows Poissonian statistics, whereas one generating pure states exhibits statistics associated with a thermal distribution [44].

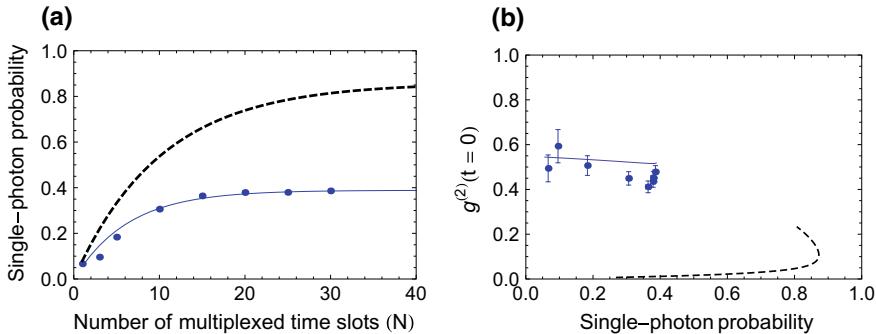


Fig. 5.4 Predicted improvement of a multiplexed heralded single-photon source. **a** Multiplexed single-photon probability versus multiplexed time slots N . **b** $g^{(2)}(t = 0)$ versus single-photon probability parameterized by p , the mean number of photon pairs. Solid lines and dots were obtained in the recent experiment [40], while dashed lines are the predicted performance with feasible improvements (heralding (95%) and detector efficiency (90%), and delay-line and switchable cavity losses (1% each))

5.2.4 Enhanced Performance

It is noteworthy to discuss further improvements to the time-multiplexed heralded single-photon source just presented. Although the source does display largely enhanced $P_M(1)$ unattainable by any non-multiplexed sources, the heralded single-photon state produced has a low purity (0.05, measured by the method proposed in [44]) because of strong spectral entanglement between signal and idler photons. To be useful for many quantum information applications utilizing multi-photon interference [15], any SPDC source must herald single photons in pure quantum states, i.e., without spectral entanglement between the signal and idler photons.⁴

Several reports [21–23] have already demonstrated pure heralded single-photon sources, utilizing perfectly or nearly degenerate photon-pair generation, where the signal-mode wavelength is identical or close to the idler-mode one. However, in the context of a time-multiplexed heralded single-photon source, since signal and idler photons have different roles (i.e., one heralds while the other is stored), it may be advantageous to generate a pair of photons at independently optimized wavelengths and bandwidths. For example, the heralding signal photon should be at a wavelength that can be detected with very high efficiency and with low latency (since the completed detection will be fed forward to store the conjugate photon). Similarly, it is beneficial that the heralded photon be generated at a low group-velocity-dispersion wavelength in the storage cavity (e.g., the group velocity dispersion at the telecom

⁴Note that the non-multiplexed single-photon probability can actually be lower for pure heralded single-photon sources than for ones generating mixed states (see Table 5.1) because of their different photon-number statistics, i.e., thermal for a single-mode SPDC source versus Poissonian for multimode states [44]; multiplexing is consequently even more important to suppress unwanted multiple-photon events.

C band (1530–1565 nm) can be an order of magnitude smaller than at 710 nm). A moderately narrower bandwidth, i.e., a longer pulse duration ($\gtrsim 1$ ps) of heralded single photons can also relax the stability requirement for the cavity length; such a requirement is important to constrain the timing jitter of a multiplexed photon, and thereby to maintain a high purity. Note that phase stabilization of the cavity length is *not* typically necessary, as the two-photon interference effects depend only on the indistinguishability of the interfering photons, and not their phases [15]. Instead, the cavity length need only be stabilized within $\sim 10\%$ of the coherence length of the heralded single photons, a much less stringent requirement; for a photon within an 1-ps-long pulse, the required level of path length precision is on the order of $10 \mu\text{m}$, which can be achieved without active stabilization.

With feasible values for heralding (95%) and detector efficiency (90%), and delay-line (1%) and switchable cavity losses (1%), we anticipate single-photon probabilities exceeding 80%, with a $g^{(2)}(t = 0)$ as low as 0.05 (see Table 5.1 and Fig. 5.4). Recent time-multiplexing implementations incorporating spectrally entangled sources have approached this, achieving $P_M(1)$ as high as 66.7% [27].

5.3 Multi-photon Source

In Sect. 5.2 we focused on time multiplexing for single-photon sources. Next we discuss multi-photon sources, where the photons are either in separate spatio-temporal modes (Sect. 5.3.1) or in a single spatio-temporal mode, such as photon-number states (Sect. 5.3.2). The former are commonly employed for quantum computing, e.g., linear optics [46] or cluster-state computing [47]; in contrast photon-number states and their related entangled version “NOON” states are known to be useful for quantum-enhanced metrology [48]. Here we present time-multiplexing methods to greatly improve multi-photon state creation efficiency.

5.3.1 *Multiple Photons in Multiple Modes*

Linear optics quantum computing (LOQC) typically requires single photons in various spatial modes interfering with each other (as we discuss in Sect. 5.5.2, recent research also raises the possibility of only using time-bin encoding [49]). One particularly interesting example is boson-sampling[50], which has the benefit that it does not require lossy active elements in the circuit; it considers the effect of an $m \times m$ multimode linear optics network whose m input modes are populated with n single photons ($m \gg n$ to avoid having multiple photons in a single mode). The network implements a unitary map on the input state. The output state will then be a superposition of the different configurations for how the n photons could have arrived in the output modes. In general, there will be $n!$ ways in which n photons could reach a specific set of n outputs (assuming each arrives at a distinct output), and the associ-

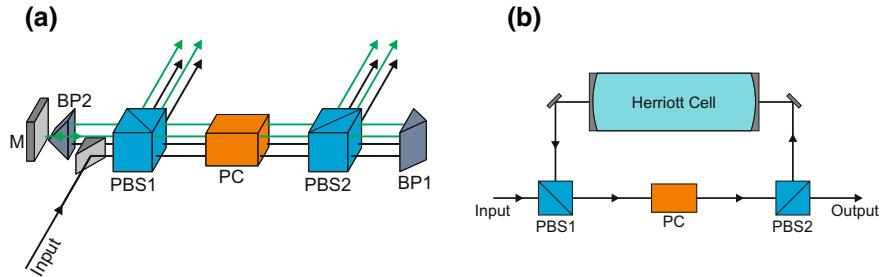


Fig. 5.5 **a** “De-Serializer” scheme to convert a stream of $n = 8$ photons in a single spatial mode into n simultaneous photons in n spatial modes. **b** Photon “compressor” which reduces the interpulse spacing of an input sequence of photons

ated amplitude will relate to an $n \times n$ matrix “permanent,” obtained from the $m \times m$ unitary matrix U describing the optics network [51]. Calculating matrix permanents is known to be a **#P**-complete problem, even harder than **NP**-complete, and the best-known algorithm requires $O(2^n n^2)$ runtime. Therefore, if boson-sampling were to be classically simulated by calculating the matrix permanents, it would require exponential classical resources. The experiment is repeated many times, each run sampling from the distribution P_S . This sampling task—reconstructing the statistical distribution at the output of the boson-sampling device—was also shown to be a computationally hard problem [52]. Thus, although there are presently no known practical applications,⁵ boson-sampling acts as an interesting proof-of-principle demonstration of “quantum supremacy”, where passive linear optics with quantum inputs can outperform classical computers.

A number of small-scale boson-sampling experiments have already been carried out [53–57]. To demonstrate larger scale boson sampling, beyond what a classical computer could solve,, one needs the capability to generate the input state with n single photons in m modes. One scheme to achieve this using a sequential (i.e., periodic) source of identical photons (as discussed in Sect. 5.2) is shown in Fig. 5.5, a simple “de-serializer” optical circuit that converts a sequence of n equally spaced photons in a single spatial mode into n simultaneous photons each in its own spatial mode. Back reflecting prisms BP1 and BP2 are used to displace the beam of photons slightly (e.g., 1 mm) in the horizontal and vertical directions, respectively, while mirror M reverses the beam direction. When all photons have entered the setup there will be a transverse 2×2 array of photons on either side of the Pockels cell. Distances PC-BP1-PC and PC-BP2-PC are adjusted to match the pulse spacing of the initial photon source; to keep the layout size manageable, the initial interpulse spacing should be

⁵Finding a computational application is challenging given the sampling aspect—if the device yields a different outcome each time it runs, how does the outcome answer a well-defined question, and how could one map it to a problem of interest?

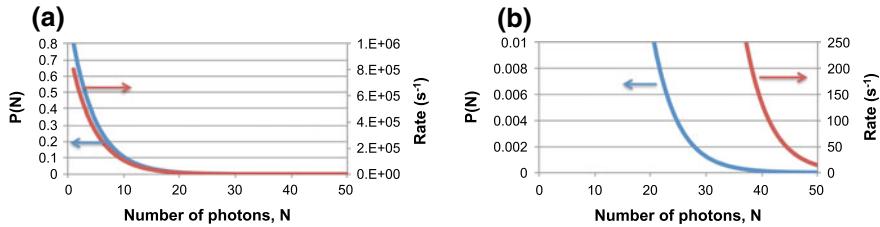


Fig. 5.6 **a** Predicted performance of our feasible source for M -sequential photon state generation, assuming a 1-MHz single-photon source rate. **b** Enlarged version at high photon number

less than 10 ns.⁶ By switching the Pockels cell, all eight photons will exit the setup at the same time through both PBS1 and PBS2, in 2×2 transverse arrays; these modes can then be coupled into optical fibers or waveguides for further processing, for example in a boson-sampling experiment. We can consider expected performance using the enhanced source described in Sect. 5.2.4. In addition, higher repetition rate (~ 1 MHz) of the time-multiplexed source is now possible due to improved PC drivers. These advances would make a significant difference for building up larger photon numbers; for example, the generation rate of a state with 8 sequential photons (which could then be converted to one with 8 simultaneous photons, each in a different spatial mode; see Fig. 5.5a) should be as high as $0.8^8 \times 1$ MHz $\sim 1.7 \times 10^5$ per second using the multiplexing techniques, while recent 8-photon experiments with non-multiplexed SPDC sources generated only ~ 1 successful event per second [25, 26]. Obviously, optimization of single-emitter-based periodic single-photon sources [11–13] would also enable those to be a competitive resource for such applications requiring multiple photons in multiple spatial modes (Fig. 5.6).

5.3.2 Fock States

For a time-multiplexed single-photon source, the goal is to create single photons with high probability by using repeated down-conversion attempts; each heralded down-conversion event adds a photon to the idler mode (both signal and idler modes are assumed to be initially in the vacuum state). We can further exploit this perspective to create multi-photon states in a single spatio-temporal mode, e.g., Fock states; such states can *not* be readily produced by single emitters [35, 36].

⁶If the initial periodic source frequency is too low, one can incorporate an optical “compressor” to reduce the interval between photons. For example, the scheme in Fig. 5.5b can convert a stream of single photons at a rate of $1/\mu\text{s}$ into a burst of $n = 8$ photons spaced by 6 ns. This is achieved by adjusting the round-trip time from PBS1 through the Herriott cell back to PBS1 to a time which is 6 ns less than the original repetition rate of the single-photon source. Thus, each time a new photon enters there will be photons leading this with a spacing of 6 ns. The Pockels cell in this setup should operate at the source repetition rate to switch each new photon into the loop. Finally, after 8 round trips the Pockels cell will be activated to emit all stored photons through PBS2.

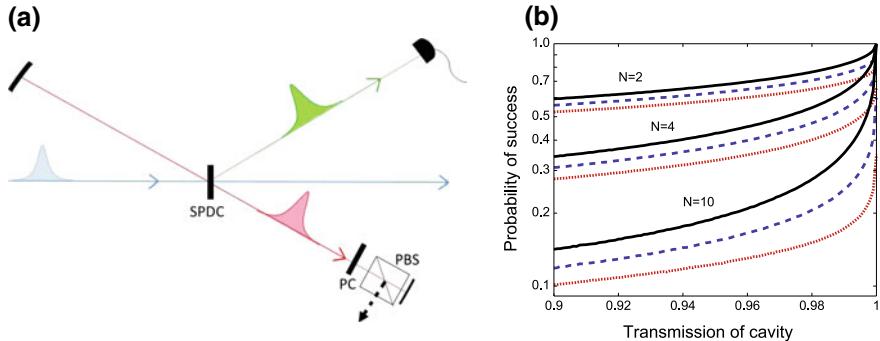


Fig. 5.7 **a** Diagram of Fock-state source. The signal photon of each pair is detected, and the idler photon is emitted into a storage cavity. When the desired number of photons is accumulated, the light is switched out by rotating polarization using a Pockels cell (PC) and the polarizing beam-splitter (PBS). **b** Theoretical performance of N -photon heralding Fock-state source. Labelled curves show several different values of N and different detector efficiencies are shown (black solid, $\eta = 1$; blue dashed, $\eta = 0.95$, red dotted, $\eta = 0.9$). Images from [41]

The setup is similar to that of the time-multiplexed single-photon source described above (Fig. 5.1), but the down-conversion crystal is instead placed *inside* the storage cavity (see Fig. 5.7a). A series of periodic pump pulses passes through the downconversion crystal. Detecting n signal photons indicates that n idler photons were emitted into the cavity (assuming a high detector efficiency). By matching the cavity round-trip time to the pump spacing, each new idler photon(s) adds to the ones already in the cavity, in precisely the same spatio-temporal mode. Thus, via this repeated single-photon addition one can “build up” the desired photon-number state; when the desired number of photons is achieved, they are switched out using the Pockels cell (PC) and PBS.

The expected performance of this method is shown in Fig. 5.7b, where a “success” is having the desired number of photons exiting the cavity, including the possibility a photon is lost from the cavity, but then replaced by a pair produced but not detected [35, 41]. This prediction assumes the ability to tune the pump amplitude for each pass through the crystal, and the ability to add multiple pairs per pulse (heralded by a photon-number resolving detector). The system performance is highly dependent on the transmission of the cavity, since many passes through the cavity are necessary to achieve high-number Fock states. Detector efficiency is also a limitation since each photon must be detected independently, but efficiencies above 90% seem adequate. We see that 10-photon states may be created with 15% probability. This could also be achieved by simply strongly pumping a non-time-multiplexed crystal and post-selecting on simultaneously detecting 10 signal photons (assuming the high photon-number resolution requirements for the heralding detector in this case). However, while both approaches require 10 pairs to be produced and detected, the time-multiplexed approach gives an additional enhancement factor of $\binom{R}{10}$, where R is the total number of pump pulses used (limited by cavity losses); in practice this

could lead to several orders of magnitude improvement in the achievable rates. However, as before the idler photons in the downconversion source need to be pure, i.e., not entangled in any way to the signal photons; otherwise, the state will be described by a tensor product of single-photon density matrices, instead of a single state, i.e., $\rho \otimes \rho \otimes \rho$ instead of $|3\rangle$.

Lastly, the photon-addition experimental setup is easily modified to instead enable one suited for nearly deterministic single-photon *subtraction*. By placing a high-transmittivity beam splitter in the path of the light field inside the recycling cavity, an annihilation operator will successfully act on the state each time a high-efficiency detector monitoring the reflected port of the beam splitter fires. The combination of controllable photon addition and subtraction could be a powerful resource in advanced quantum state engineering.

5.4 Digital Quantum Memory

A variable-time delay photon quantum memory has many applications in quantum computation and communication. For instance, it provides precise synchronization of qubits in quantum repeater architectures, and is completely essential for large-scale linear optics quantum computing, with probabilistic gates that may need to be attempted multiple times without losing synchronization with the rest of the computer. An optical system can also be a mechanism to convert heralded photons into on-demand photons, while leaving the state itself unchanged (essentially an identity quantum gate). Though this paper will focus on all-optical quantum memories, there is a great deal of research in atomic, ensemble, and solid-state quantum memories [58–62]. One typical limitation of these is the effective bandwidth of the memory; bandwidths below ~ 10 MHz may be undesirable, as they require separating the photons by < 100 ns to prevent crosstalk.

5.4.1 “Digital” Delay Line

There are many approaches to developing a quantum memory, the simplest of which is to store light in an optical delay line, such as a fiber or a high-Q cavity. However, to be useful in a synchronization capacity, the duration of the delay must be adjustable in real time. For example, a single loop with a switching element can be used to delay any incident photon by any integer number of roundtrips (leftmost loop in Fig. 5.8a). The problem is that while the mirrors can be made nearly lossless (less than 0.01%), the active switch elements (e.g., a Pockels cell and polarizing beam splitter) typically display losses and crosstalk of $\sim 1\%$, thereby limiting the delay to only ~ 50 cycles; the problem is much worse for integrated optics implementations, where 10% loss is considered low. One method to extend the useful range of operation is to link multiple delays, each with a different delay timescale (entire system in Fig. 5.8a). For

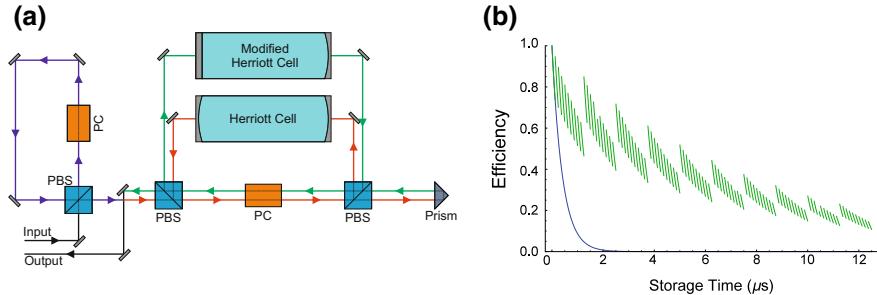


Fig. 5.8 **a** All-optical quantum memory, combining 3 delay lines with different effective storage times. **b** Predicted efficiency versus storage time. The green curve represents the net transmission of 3 different delay cavities and 2 Pockel cells (the entire system in **(a)**); for comparison, the blue curve represents the transmission of a simple 12.5-ns delay line with a Pockel cell (only the leftmost loop in **(a)**). Here we assume mirror reflectivities of 99.9% and a Pockel Cell Transmission of 97.5%. The use of Herriott cells in the setup allows for significantly improved overall efficiency, especially at longer time delays

example, we have created a “digital delay” by using multiple optical cavities with delay times of 12.5 ns, 125 ns, and 1.25 μ s. The times are chosen to match the repetition rate of a typical mode-locked Ti:Sapph laser. To achieve the two larger cavities’ delay times, we use a Herriott cell [63] with 19 reflections and modified Herriott cell [43] with 170 reflections. The theoretical efficiency versus storage time for such a system is shown in Fig. 5.8b; the predicted average transmission over delays from 1 to 999 cycles is 33%, much higher than the average if a single switchable loop were used (0.00004%), assuming the same 1% Pockels cell and PBS losses. While such a system could be entirely implemented using fiber-optic delays [64], a free-space memory benefits from the availability of broad bandwidth, high-reflectivity dielectric coatings and relatively low-loss electro optic switches; a theoretical Pockel cell bandwidth is ± 80 nm, with a transmission of 97%. Assuming an operating wavelength of 690 nm, this gives us an estimate of 5×10^8 (40 THz \times 12.5 μ s) for our time-bandwidth product, several orders of magnitude higher than typical memories based on atomic ensembles [59] or solid-state approaches [65].

It is desirable to be able to store optical quantum information carried in various degrees of freedom, i.e., polarization, timing, spatial mode, etc. For example, as long as the bin spacing is less than the minimum loop size, the above system would be able to faithfully store time-bin qubits as discussed in Sect. 5.5.2. If the memory is implemented using bulk optics (as opposed to single-mode fiber delay lines) one may also be able to store spatial-mode quantum states, e.g., various orbital angular momentum states. Due to the dependence of the proposed optical storage on the manipulation of polarization, however, in order to store polarization qubits we need a method to convert polarization encoding into time-bin encoding; in effect, we need a polarization-to-time-bin qubit transducer. For example, as shown in Fig. 5.9a, we can use two polarizing beamsplitters and a Pockels cell to convert a polarization qubit in a single time bin into a single-polarization two-time-bin qubit [66], which

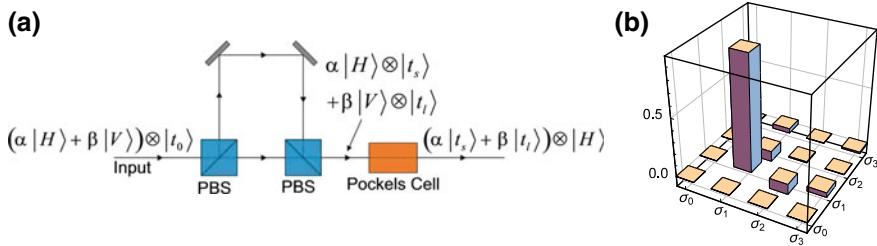


Fig. 5.9 **a** A polarization-to-time-bin transducer. Polarizing beam splitters are used to convert two polarization states in a single time bin to two separated time bins. A fast electro-optic switch then rotates the polarization of one of the pulses to match the other, yielding a single-polarization time-bin qubit. **b** The measured superoperator describing the quantum process of the polarization to time-bin transducer. This matrix represents a bit-flip operation, as expected. The fidelity with χ is 0.978 ± 0.001

can then be directed into the above digital delay, for example. After a photon has been stored for the desired time, a similar system can convert the timing information back into polarization. We have implemented such a transducer, at present using the same optical elements to convert polarization into time bin, and to implement the conversion back into polarization (by directing the time-bin qubit back through the transducer). We tested the six input states ($|H\rangle, |V\rangle, |D\rangle, |A\rangle, |L\rangle$, and $|R\rangle$), using a polarizing beam splitter and waveplate to measure each state in three mutually unbiased polarization bases; the state was then reconstructed via the maximum likelihood method of quantum state tomography, and purity and fidelity were computed from the reconstructed density matrix [67, 68]. After converting to time-bin qubits and then back to polarization qubits, we observed fidelities greater than 0.973 in all cases ($F_{\text{average}} = 0.998 \pm 0.009$, where the uncertainty is determined from 10 independent trials) and average purity $P_{\text{average}} = 0.996 \pm 0.004$. The double-pass through the transducer performs a bit-flip operation (changing horizontal polarization to vertical and vice versa) in this test; this bit-flip operation can be characterized by a superoperator χ :

$$\chi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (5.3)$$

Quantum process tomography was performed to measure the superoperator of the transducer and compare it to χ [53]. The result is shown in Fig. 5.9b; the fidelity of the superoperator with χ is 0.978 ± 0.001 .

Finally, the memory could also enable efficient storage of hyperentangled photons, with information stored in all these degrees of freedom simultaneously. While the first hyperentangled storage systems [69] only achieved 5% efficiency, the estimated net average transmission with a bulk optics digital delay cell exceeds 50% for delay times up to 6.25 μs .

5.4.2 Memory Cavity and Shutter Cavity

Another form of all-optical quantum memory uses two cavities; the photons are both created and stored inside one cavity and released by dynamically tuning the other cavity [70, 71]. A simplified schematic of the setup is shown in Fig. 5.10a. Inside a nondegenerate optical parametric oscillator (NOPO), signal and idler photons are probabilistically created and simultaneously appear in different longitudinal modes. A shutter cavity (SC) is placed at the exit of the NOPO; the resonance frequency of the SC can be shifted using an electro-optic modulator (EOM). A photon inside the NOPO passes through the SC only when in resonance (otherwise, it stays within the NOPO) (see Fig. 5.10b). The NOPO can therefore be labelled as a memory cavity (MC).

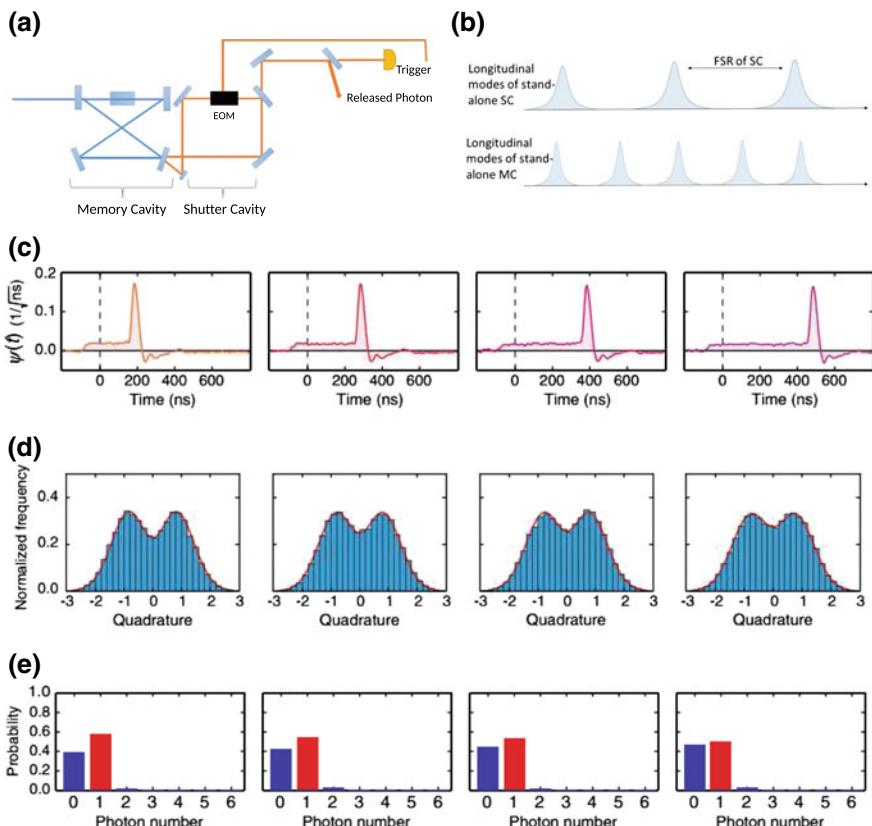


Fig. 5.10 Schematic diagrams of storage and release. **a** Experimental setup and beam paths. **b** Conceptual diagram in the frequency domain. **c** Estimated wave-packet envelopes of the released photons. The idler-photon detection event corresponds to 0 ns. **d** Sectional side views of the Wigner functions, cutting through the phase-space origin. **e** Photon-number distributions. From [71]

Figure 5.10c, d, and e show experimental analysis of storage time at 0, (intrinsic delay, leftmost column) 100, 200, and 300 ns. The temporal shape of the wave-packet envelope $\psi(t)$ is estimated from the raw data from homodyne detection [71], as is the Wigner function. After the shutter opens at about 150 ns, the wave-packet amplitude suddenly increases and then rapidly goes to zero—this represents the emission of the photon. Though total photon losses of 40% were observed, only about 15% of photon losses are estimated to be coming from the shutter cavity. Transmissions of 54.6%, 53.1%, and 49.7% are obtained for 100, 200, and 300 ns, respectively. The lifetime of the photon storage is of the order of 1 μs ; longer storage renders the Wigner function positive, but the single photons may still be useful, owing to their relatively small multi-photon components; see Fig. 5.10e for an idea of low photon numbers. Moreover, decreased intracavity losses should extend the useful storage times.

The setup in Fig. 5.10a has been used to synchronize the arrival times of two photons, preserving their purity [70]. Controlled storage times up to 1.8 μs were achieved with purities sufficiently high for a negative Wigner function.

5.5 Applications

Finally, we discuss several applications of time-bin and time-multiplexing methods. This is not intended to be a comprehensive list but rather representative of the current state of interesting research.

5.5.1 Time-Multiplexed Single-Photon Detectors

A limitation of typical photon-counting avalanche photodiodes (APDs) is that they only produce a single output pulse, with uniform height, regardless of the number of incident photons, i.e., there is no way to distinguish if a pulse contains 1 or $n > 1$ photons. By incorporating time-multiplexing techniques, non-photon-number resolving detectors can be enhanced to display modest photon-number resolution [72, 73]. As indicated in Fig. 5.11a, time multiplexing divides the incident pulse into N separate pulses of approximately equal amplitude that are displaced by time interval δT , larger than the detector dead time. For $N \gg n$, the divided pulses have a negligible probability of containing two or more photons in a single time window, thus overcoming the fact that an APD can only produce a single output pulse. A theoretical analysis of the photon detector has been performed [72]—it would seem that a large limitation of the device rests on high detection efficiency and near lossless fiber, though these traits would also be desirable for *non-time-multiplexed* single-photon detectors. Figure 5.11b shows probability distributions for various photon numbers and detection efficiencies; even with 100%-efficient detectors, eight temporal modes are needed to effectively discriminate the photon numbers for coherent states with

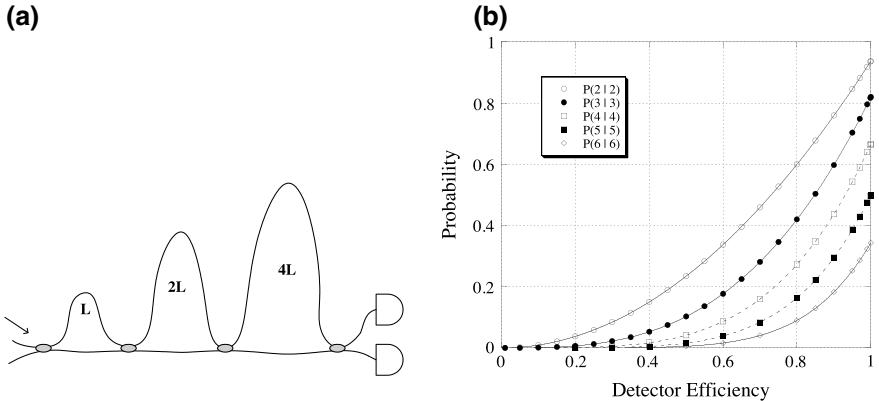


Fig. 5.11 **a** Implementation of a time-multiplexed detector. An incident pulse is divided into N timeslots, separated by more than the detector dead time using a series of fiber couplers and optical delay lines. **b** Calculated probability $P(n|n)$ of detecting all n incident photons as a function of the effective single-photon detection efficiency η . Taken from [72]

mean photon number as high as 1.5 [73]. In addition, the use of time multiplexing lengthens the effective detector response time, which may be a disadvantage in certain high-rate applications.

5.5.2 Time-Bin Qubits

Although formally any two-state representation of quantum information is equivalent to any other, in practice the explicit physical implementations can have various advantages or disadvantages. For example, polarization has long been the most commonly used degree of freedom for encoding quantum information in photons, due to the ease in manipulating and measuring this degree of freedom. However, the use of time-bin encoding offers several advantages. First, the technique is very robust, in the sense that as soon as one is able to demonstrate a single-photon emitter (e.g., a quantum dot, single ion, nitrogen-vacancy (NV) center, etc.,) in principle one can produce a time-bin quantum bit simply by driving the emitter with a pair of coherent optical pulses separated by an interval less than the decoherence time of the system. This has been done to demonstrate time-bin entangled photons from downconversion [74] and also from quantum dots [75, 76].

Another advantage of time-bin encoding is that these states tend to be more robust in propagation through optical fibers (unlike polarization qubits) and also through atmospheric turbulence (unlike spatial-mode encoding, e.g., using orbital angular momentum states [77–79]).⁷ The main reason is that fast detectors allow one to use

⁷One potential disadvantage for time-bin encoding when using a free-space communication channel is that the unbalanced interferometers used to analyze the states are usually only reliable for a single

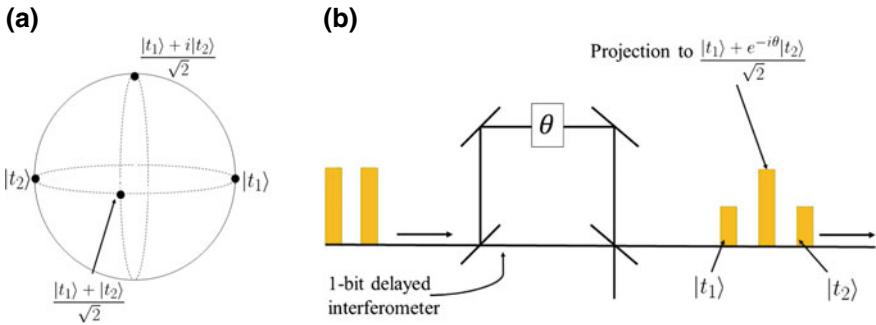


Fig. 5.12 **a** The equivalent Bloch sphere for time-bin qubit encoding. The early and late time bins are represented by the poles of the sphere, while equal-amplitude superpositions reside on the equator. **b** Method to (probabilistically) project onto various time-bin qubit states. Photons detected in the extremal pulses correspond to measurements in the two time bins directly, while photons detected in the middle pulse correspond to equal-amplitude superposition measurements

time bins that are separated by less than a few nanoseconds, and this is not enough time for the channel fiber or atmosphere to act differently on adjacent time bins.

Moreover, the timing degree of freedom, being essentially continuous, can enable storage of much more interesting quantum states. For example, instead of driving the emitter with two temporally separated pump pulses, one can use three, four, or many more. When this is done using spontaneous parametric downconversion, one can easily produce photon entangled pairs which are entangled in a 1000-dimensional Hilbert space [81]. This in turn can lead to the ability to transmit multiple quantum bits per photon in an advanced quantum cryptography protocol [81–83] (though one does need to go to some lengths to verify the presence of an eavesdropper, either using a mutually unbiased basis to the time bins [84–86] or in some cases by using another degree of freedom altogether, e.g., polarization [81, 87]); this can also be accomplished using weak coherent states and pulse-position modulation [88]. Returning for simplicity to the qubit state, techniques for performing complete quantum state tomography have been developed; an arbitrary quantum bit of this sort can still be represented by a location on the Bloch sphere, with the individual time-bin states on two antipodal points, and various superpositions in other locations on the sphere; see Fig. 5.12a.

Figure 5.12b shows how one can probabilistically project onto the canonical time-bin states needed for quantum state tomography, using an unbalanced interferometer whose path length difference matches the time-bin spacing [89]. Photons in the early (late) time bin which take the short (long) path end up in the first (third) pulse; the

spatial mode, i.e., multiple spatial modes will each have a different path length imbalance, degrading the overall system performance unless one accepts the loss of single-mode filtering; however, by including “4f”-imaging optics in both arms—essentially imaging the first beamsplitter onto the second—one can achieve a path imbalance that is independent of the incident beam tilt [80]. We have realized such a system, and demonstrated multi-mode visibilities above 93%, for input tilt angles up to 150 μrad.

middle pulse measures equal superpositions of the two time bins (with a relative phase θ determined by the phase in the interferometer). In order to make this a deterministic scheme, one would need a fast switchable first beam splitter, e.g., effectively forcing photons in the early (late) time bin to take the long (short) path, as was discussed in Sect. 5.4.1. In any event, the minimum time-bin separation must be greater than the detector resolution; with existing superconducting nanowire detectors this can be as low as 30 ps [90]. Alternatively, using nonlinear frequency conversion techniques with chirped pump pulses, one can transform arbitrary time-bin states, including superpositions of time bins, into spectrally resolvable frequency states; via the process of sum-frequency generation, photons at slightly different times convert with different spectral components of the chirped pump pulse, yielding spectrally resolvable frequencies [91]. The advantage is that time-bin qubits with much smaller spacings (~ 1 ps) can be accurately measured, even with detectors with nanosecond-scale temporal resolution (though at the cost of additional experimental complexity).

Methods to deterministically prepare arbitrary time-bin qubits (and higher dimensional qudits) have been proposed. One option is to first convert the time-bin encoding to polarization (as discussed in Sect. 5.4.1), and then use standard birefringent optics to create the desired state, before converting it back into time-bin encoding [49]; an alternative proposal (which need not disturb the polarization state) uses a series of interconnected interferometers, with fast controllable phase modulators in each [92].

One recent application of time-bin encoding is the realization that one could in principle implement linear optics quantum computing operations using *only* various time-bin qubits in a single spatial mode; such a scheme was used to demonstrate a post-selected Controlled-PHASE gate with 84% gate fidelity [49].

5.5.3 Random Walks

The quantum walk model describes the coherent propagation of quantum particles under stepwise unitary evolutions. A quantum walker is described by a quantum mechanical wave function which the evolution spreads over different positions, the system achieving a coherent superposition of all paths. The resulting quantum interference effects drastically change the final probability distribution; for instance, one can observe a net spreading that varies linearly with the step size, in contrast to the usual square-root dependence seen in classical random walks, e.g., diffusion. Quantum walks have many applications in quantum computation [94, 95], and many groups have explored various optical implementations using walks between different spatial modes, e.g., adjacent waveguides [96, 97]. Discrete-time quantum walks can also be realized experimentally using time-multiplexing; photons travel a longer or shorter path depending on their polarization state—this corresponds to a step to the left or the right. The different arrival times then correspond to the different spatial positions [93]. Figure 5.13 shows an experimental implementation of a 1D quantum walk, where a laser pulse travels two different paths in a fiber network, correspond-

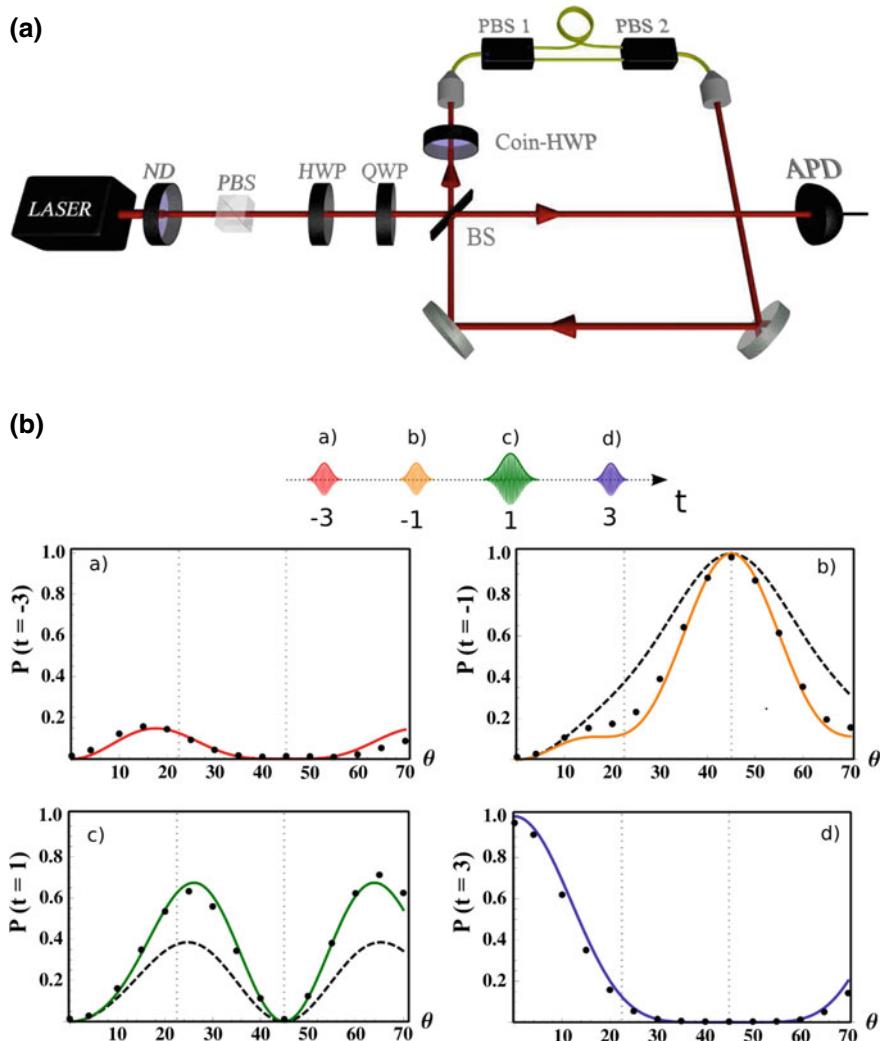


Fig. 5.13 **a** Sketch of the setup to demonstrate a time-based 1-D quantum walk. Using neutral density filters (ND), a laser field is attenuated to the single-photon level and coupled into the loop using a beam splitter (BS). A half wave plate in the loop serves to realize the “coin” operation of the quantum walk protocol, making polarization the “coin” space in which the photon is rotating. **b** A demonstration of how the probability distribution after three steps changes as a function of the angle of the half-wave plate; the photon is initially prepared with horizontal polarization. Dots: Measured quantum walk (error bars are smaller than used symbols). Solid line: theoretical model for quantum walk; dashed line: classical random walk. Taken from [93]

ing to the two possible directions in a quantum walk on a line. The different arrival times then correspond to the different spatial positions [93]. This technique can be generalized to simulate quantum walks in multiple directions; for example, [98] uses four different paths at each node in a fiber network. With proper adjustments, such systems can simulate the dynamics of quantum particles traversing a homogeneous system influenced by a tunable source of decoherence. This allows for a controlled transition between the coherent quantum mechanical dynamics to an incoherent evolution where the distribution is determined by classical randomness.

5.5.4 *Entanglement Swapping*

While many quantum communication protocols, e.g., quantum cryptography [99], superdense coding [100–102], quantum fingerprinting [103], quantum teleportation [104], etc., have now been demonstrated over relatively short links to achieve global-scale operation will require the realization of quantum repeater networks [105, 106], in order to distribute entanglement in the presence of channel loss, e.g., absorption in optical fibers [107]. At the heart of every quantum repeater [108] is the technique of “entanglement swapping,” by which entanglement can be transferred onto two systems that never interacted with each other [109, 110]. For example, if we have two independent entangled photons pairs (A_1, A_2) and (B_1, B_2), a projective Bell-state measurement between A_1 and B_1 will collapse the remaining two photons (A_2 and B_2) into an entangled state.

In a quantum repeater network, different parties may be separated by multiple nodes where entanglement swapping occurs. If one successfully achieves Bell-state projection between the photon pairs at every node, then entanglement is generated between the first and last nodes. However, Bell-state measurement requires precise overlap of the photons, which in practice is difficult to achieve if the entangled photon pairs are generated in different sources. While many groups have demonstrated entanglement swapping [108] and even multi-stage entanglement swapping [112], the entangled photons in those experiments were created and measured at the same time.

To overcome this issue, a quantum memory system is essential for storing the entangled photons until the system is ready for entanglement swapping. For proof of concept, Megidish et al. showed how a time-multiplexed downconversion source mimics a quantum repeater setup with a memory system [111]. In this setup, entanglement is generated between time-multiplexed downconversion events. A pulsed laser is used to generate consecutive photon pairs separated by the laser period time τ (see Fig. 5.14). The first photon (1) from the first downconversion pair is measured while the second photon (2) is delayed by τ , using highly reflective dielectric mirrors to create a free-space delay line with an overall transmission of 90% (see Fig. 5.14b). After the second pair of entangled photons is generated (3 and 4), Bell

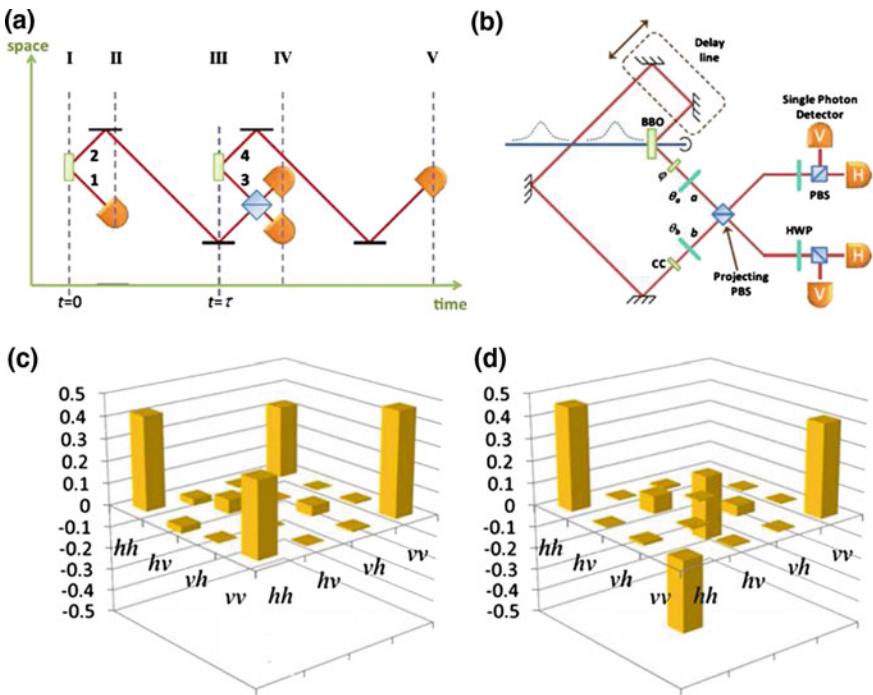


Fig. 5.14 Description for between timelike separated quantum systems, Photons 1 and 4. **a** Timeline diagram. (I) Birth of first photon pair (1 and 2), (II) detection of photon 1, (III) birth of second photon pair (3 and 4), (IV) Bell state measurement of photon (2 and 3), (V) detection of last photon. **b** Experimental setup. A BBO crystal is used to produce degenerate downconversion pairs at 780 nm, and compensating crystals (CC) are placed in each arm to correct for temporal walk-offs and for control over the phase (ϕ) of the state. Finally, half-wave plates (denoted by θ in the figure) are also inserted to analyze the photons in a rotated basis. A free-space delay line is used to introduce an optical delay on one of the photons from the first downconversion pair. **c**, **d** show real parts of the density matrices of the first and last photons. **c** The two middle photons are projected onto the $|\psi^+\rangle$ state. **d** The two middle photons are projected onto the $|\psi^-\rangle$ state. Images adapted from [111]

projection occurs between the nondelayed photon of the second pair (3) and the previously delayed photon (2). The last photon (4) is delayed by τ before measurement. The experiment verified entanglement between photons 1 and 4 with a fidelity of $77 \pm 1\%$ [112]. The scenario described above is similar to that of a practical quantum repeater setup: when one of the entangled photons reaches a node, it will be stored in a quantum memory, waiting for the second entangled photon to arrive. During this delay, other entangled photons can already be used or sent to other nodes.

References

1. T. Pittman, It's a good time for time-bin qubits. *Physics* **6**, 110 (2013)
2. T.M. Graham, J.T. Barreiro, M. Mohseni, P.G. Kwiat, Hyperentanglement-enabled direct characterization of quantum dynamics. *Phys. Rev. Lett.* **110**, 060404 (2013)
3. M.D. Eisaman, J. Fan, A. Migdall, S.V. Polyakov, Invited review article: single-photon sources and detectors. *Rev. Sci. Instrum.* **82**, 071101 (2011)
4. J. McKeever, A. Boca, A.D. Boozer, R. Miller, J.R. Buck, Deterministic generation of single photons from one atom trapped in a cavity. *Science* **303**, 1992–1994 (2004)
5. M. Hijlkema, B. Weber, H.P. Specht et al., A single-photon server with just one atom. *Nat. Phys.* **3**, 253–255 (2007)
6. M. Keller, B. Lange, K. Hayasaka, W. Lange, H. Walther, Continuous generation of single photons with controlled waveform in an ion-trap cavity system. *Nature* **431**, 1075–1078 (2004)
7. P. Michler, A. Kiraz, C. Becher et al., A quantum dot single-photon turnstile device. *Science* **290**, 2282–2285 (2000)
8. A.J. Bennett, D.C. Unitt, P. Atkinson et al., High performance single photon sources from photolithographically defined pillar microcavities. *Opt. Express* **13**, 50–55 (2005)
9. A. Beveratos, R. Brouri, T. Gacoin et al., Single photon quantum cryptography. *Phys. Rev. Lett.* **89**, 187901 (2002)
10. N. Mizuochi, T. Makino, H. Kato et al., Electrically driven single-photon source at room temperature in diamond. *Nat. Photon.* **6**, 299–303 (2012)
11. C. Santori, D. Fattal, J. Vučković et al., Indistinguishable photons from a single-photon device. *Nature* **419**, 594–597 (2002)
12. Y.-J. Wei, Y.-M. He, M.-C. Chen et al., Deterministic and robust generation of single photons from a single quantum dot with 99.5% indistinguishability using adiabatic rapid passage. *Nano Lett.* **14**, 6515–6519 (2014)
13. A.K. Nowak, S.L. Portalupi, V. Giesz et al., Deterministic and electrically tunable bright single-photon source. *Nat. Commun.* **5**, 3240 (2014)
14. X. Ding, Y. He, Z.-C. Duan et al., On-demand single photons with high extraction efficiency and near-unity indistinguishability from a resonantly driven quantum dot in a micropillar. *Phys. Rev. Lett.* **116**, 020401 (2016)
15. C.K. Hong, Z.Y. Ou, L. Mandel, Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**, 2044–2046 (1987)
16. J.-W. Pan, Z.-B. Chen, C.-Y. Lu et al., Multiphoton entanglement and interferometry. *Rev. Mod. Phys.* **84**, 777–838 (2012)
17. C.K. Hong, L. Mandel, Experimental realization of a localized one-photon state. *Phys. Rev. Lett.* **56**, 58–60 (1986)
18. B.G. Christensen, K.T. McCusker, J.B. Altepeter et al., Detection-loophole-free test of quantum nonlocality, and applications. *Phys. Rev. Lett.* **111**, 130406 (2013)
19. M.D.C. Pereira, F.E.B. Becerra, B.L. Glebov et al., Demonstrating highly symmetric single-mode, single-photon heralding efficiency in spontaneous parametric downconversion. *Opt. Lett.* **38**, 1609 (2013)
20. M. Giustina, A. Mech, S. Ramelow et al., Bell violation using entangled photons without the fair-sampling assumption. *Nature* **497**, 227–230 (2013)
21. P.J. Mosley, J.S. Lundeen, B.J. Smith, I.A. Walmsley, Conditional preparation of single photons using parametric downconversion: a recipe for purity. *New J. Phys.* **10**, 093011 (2008)
22. P.G. Evans, R.S. Bennink, W.P. Grice et al., Bright source of spectrally uncorrelated polarization-entangled photons with nearly single-mode emission. *Phys. Rev. Lett.* **105**, 253601 (2010)
23. J.B. Spring, P.S. Salter, B.J. Metcalf et al., On-chip low loss heralded source of pure single photons. *Opt. Express* **21**, 13522–13532 (2013)

24. R. Krischek, W. Wieczorek, A. Ozawa et al., Ultraviolet enhancement cavity for ultrafast nonlinear optics and high-rate multiphoton entanglement experiments. *Nat. Photonics* **4**, 170–173 (2010)
25. Y.-F. Huang, B.-H. Liu, L. Peng et al., Experimental generation of an eight-photon Greenberger-Horne-Zeilinger state. *Nat. Commun.* **2**, 546 (2011)
26. X.-C. Yao, T.-X. Wang, P. Xu et al., Observation of eight-photon entanglement. *Nat. Photonics* **6**, 225–228 (2012)
27. F. Kaneda, P.G. Kwiat, High-efficiency single-photon generation via large-scale active time multiplexing. [arxiv:1803.04801v1](https://arxiv.org/abs/1803.04801v1) (2018)
28. A.L. Migdall, D. Branning, S. Castelletto, Tailoring single-photon and multiphoton probabilities of a single-photon on-demand source. *Phys. Rev. A* **66**, 053805 (2002)
29. X. Ma, S. Zotter, J. Kofler et al., Experimental generation of single photons via active multiplexing. *Phys. Rev. A* **83**, 043814 (2011)
30. M.J. Collins, C. Xiong, I.H. Rey et al., Integrated spatial multiplexing of heralded single-photon sources. *Nat. Commun.* **4**, 2582 (2013)
31. R.J.A. Francis-Jones, P.J. Mosley, Exploring the limits of multiplexed photon-pair sources for the preparation of pure single-photon states. [arXiv:1409.1394](https://arxiv.org/abs/1409.1394) (2014)
32. D. Bonneau, G.J. Mendoza, J.L. O'Brien, M.G. Thompson, Effect of loss on multiplexed single-photon sources. *New J. Phys.* **17**, 043057 (2015)
33. T. Pittman, B. Jacobs, J. Franson, Single photons on pseudodemand from stored parametric down-conversion. *Phys. Rev. A* **66**, 042303 (2002)
34. E. Jeffrey, N.A. Peters, P.G. Kwiat, Towards a periodic deterministic source of arbitrary single-photon states. *New J. Phys.* **6**, 100 (2004)
35. K. McCusker, P.G. Kwiat, Efficient optical quantum state engineering. *Phys. Rev. Lett.* **103**, 163602 (2009)
36. B. Glebov, J. Fan, A. Migdall, Deterministic generation of single photons via multiplexing repetitive parametric downconversions. *Appl. Phys. Lett.* **103**, 031115 (2013)
37. J. Mower, D. Englund, Efficient generation of single and entangled photons on a silicon photonic integrated chip. *Phys. Rev. A* **84**, 052326 (2011)
38. G.J. Mendoza, R. Santagati, J. Munns et al., Active temporal multiplexing of photons. *Optica* **3**, 127–132 (2016)
39. C.T. Schmiegelow, M.A. Larotonda, Multiplexing photons with a binary division strategy. *Appl. Phys. B* **74**, 902 (2013)
40. F. Kaneda, B.G. Christensen, J.J. Wong et al., Time-multiplexed heralded single-photon source. *Optica* **2**, 1010–1013 (2015)
41. K.T. McCusker, Efficient quantum optical state engineering and applications. Ph.D. thesis (University of Illinois at Urbana-Champaign, 2012)
42. A.I. Lvovsky, B.C. Sanders, W. Tittel, Optical quantum memory. *Nat. Photonics* **3**, 706–714 (2009)
43. C. Robert, Simple, stable, and compact multiple-reflection optical cell for very long optical paths. *Appl. Opt.* **46**, 5408–5418 (2007)
44. A. Christ, K. Laiho, A. Eckstein et al., Probing multimode squeezing with correlation functions. *New J. Phys.* **13**, 033027 (2011)
45. R. Loudon, *The Quantum Theory of Light*, vol. 3, (Oxford University Press, 2000)
46. C.R. Myers, R. Laflamme, Linear optics quantum computation: an overview. [arXiv:0512104](https://arxiv.org/abs/0512104) (2005)
47. M.A. Nielsen, Cluster-state quantum computation. *Rep. Math. Phys.* **47**, 147–161 (2006)
48. P. Kok, S.L. Braunstein, J.P. Dowling, Quantum lithography, entanglement and Heisenberg-limited parameter estimation. *J. Opt. B* **13**, 033027 (2011)
49. P.C. Humphreys, B.J. Metcalf, J.B. Spring et al., Linear optical quantum computing in a single spatial mode. *Phys. Rev. Lett.* **111**, 150501 (2013)
50. S. Aaronson, A. Arkhipov, The computation complexity of linear optics, in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing* (2011), pp. 333–342

51. L. Valiant, The complexity of computing the permanent. *Theor. Comput. Sci.* **8**, 189–201 (1979)
52. J.P. Buhler, H.W. Lenstra Jr., C. Pomerance, Factoring integers with the number field sieve, in *The Development of the Number Field Sieve* (1993), pp. 50–94
53. M.A. Broome, A. Fedrizzi, S. Rahimi-Keshari et al., Photonic Boson sampling in a tunable circuit. *Science* **339**, 794–798 (2013)
54. J.B. Spring, B.J. Metcalf, P.C. Humphreys et al., Boson sampling on a photonic chip. *Science* **339**, 798–801 (2013)
55. M. Tillmann, B. Dakić, R. Heilmann et al., Experimental Boson sampling. *Nat. Photonics* **7**, 540–544 (2013)
56. N. Spagnolo, C. Vitelli, M. Bentivegna et al., Experimental validation of photonic Boson sampling. *Nat. Photonics* **8**, 615–620 (2014)
57. M. Bentivegna, N. Spagnolo, C. Vitelli et al., Experimental scattershot Boson sampling. *Sci. Adv.* **1**, e1400255 (2015)
58. A. Dantan, J. Ciklinski, M. Pinard, Ph Grangier, Dynamics of a pulsed continuous-variable quantum memory. *Phys. Rev. A* **73**, 032338 (2006)
59. J. Jin, E. Saglamyurek, M. Ií, G. Puigibert et al., Telecom-wavelength atomic quantum memory in optical fiber for heralded polarization qubits. *Phys. Rev. Lett.* **115**, 140501 (2015)
60. J. Appel, E. Figueira, D. Korystove et al., Quantum memory for squeezed light. *Phys. Rev. Lett.* **100**, 093602 (2008)
61. M. Gündoðan, P.M. Ledingham, K. Kutluer et al., Solid state spin-wave quantum memory for time-bin qubits. *Phys. Rev. Lett.* **114**, 230501 (2015)
62. V. Parigi, V. D'Ambrosio, C. Arnold et al., Storage and retrieval of vector beams of light in a multiple-degree-of-freedom quantum memory. *Nat. Commun.* **6**, 7706 (2015)
63. A. Sennaroglu, J. Fujimoto, Design criteria for Herriott-type multi-pass cavities for ultrashort pulse lasers. *Opt. Express* **11**, 1106–1113 (2003)
64. G.G. Ball, W.H. Glenn, W.W. Morey, Programmable fiber optic delay line. *IEEE Photonics Technol. Lett.* **6**, 741–743 (1994)
65. E. Saglamyurek, A quantum memory for orbital angular momentum photonic qubits. *Nat. Photonics* **8**, 234–238 (2014)
66. Y. Soudagar, F. Bussières, G. Berlin, S. Lacroix, J. Fernandez, Cluster-state quantum computing in optical fibers. *J. Opt. Soc. Am. B* **24**, 226–230 (2007)
67. J. Altepeter, E. Jeffrey, P. Kwiat, Photonic state tomography. *Adv. At. Mol. Opt. Phys.* **52**, 105–159 (2005)
68. D.V. James, P. Kwiat, W. Munro, A. White, Measurement of qubits. *Phys. Rev. A* **64**, 052312 (2001)
69. A. Tiranov et al., Storage of hyperentanglement in a solid-state quantum memory. *Optica* **2**, 287–297 (2015)
70. K. Makino, Y. Hashimoto, J.-I. Yoshikawa, H. Ohdan, T. Toyama, P. vanLoock, A. Furusawa, Synchronization of optical photons for quantum information processing. *Sci. Adv.* **2**, e150177 (2016)
71. J.-I. Yoshikawa, K. Makino, S. Kurata et al., Creation, storage, and on-demand release of optical quantum states with a negative wigner function. *Phys. Rev. X* **3**, 041028 (2013)
72. M.J. Fitch, B.C. Jacobs, T.B. Pittman, J.D. Franson, Photon-number resolution using time-multiplexed single-photon detectors. *Phys. Rev. A* **68**, 043814 (2003)
73. D. Achilles, C. Silberhorn, C. Śliwa, K. Banaszek, I.A. Walmsley, Fiber-assisted detection with photon number resolution. *Opt. Lett.* **28**, 2387–2389 (2003)
74. J. Brendel, N. Gisin, W. Tittel, H. Zbinden, Pulsed energy-time entangled twin-photon source for quantum communication. *Phys. Rev. Lett.* **82**, 2594–2597 (1999)
75. H. Jayakumar, A. Predojević, T. Kauten, T. Huber, G.S. Solomon, G. Weihs, Time-bin entangled photons from a quantum dot. *Nat. Commun.* **5**, 4251 (2014)
76. M.A.M. Versteegh, M.E. Reimer, A.A. vanden Berg et al., Single pairs of time-bin-entangled photons. *Phys. Rev. A* **92**, 033802 (2015)

77. S. Etcheverry, G. Cañas, E.S. Gómez et al., Quantum key distribution session with 16-dimensional photonic states. *Sci. Rep.* **3**, 02316 (2013)
78. M. Malik, M. O'Sullivan, B. Rodenburg et al., Influence of atmospheric turbulence on optical communications using orbital angular momentum for encoding. *Opt. Exp.* **20**, 13195–13200 (2012)
79. M. Mafu, A. Dudley, S. Goyal et al., Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases. *Phys. Rev. A* **88**, 032305 (2013)
80. J. Jin, S. Agne, J. -P. Bourgois, Y. Zhang, T. Jennewein et al., Demonstration of analyzers for multimode photonic time-bin qubits. *Phys. Rev. A* **97**, 043847 (2018)
81. B. Christensen, K. McCusker, D. Gauthier, D. Kumor, V. Chandar, P. Kwiat, Higher-dimensional quantum cryptography. *OSA Tech. Dig.* **3**, 2316 (2013)
82. T. Zhong, H. Zhou, R.D. Horansky et al., Photon-efficient quantum key distribution using time-energy entanglement with high-dimensional encoding. *New J. Phys.* **17**, 022002 (2015)
83. D. Bunandar, Z. Zhang, J. Shapiro, D. Englund, Practical high-dimensional quantum key distribution with decoy states. *Phys. Rev. A* **91**, 022336 (2015)
84. T. Brougham, S. Barnett, K. McCusker, P.G. Kwiat, D. Gauthier, Security of high-dimensional quantum key distribution protocols using Franson interferometers. *J. Phys. B: At. Mol. Opt. Phys.* **46**, 104010 (2013)
85. T. Brougham, S. Barnett, Mutually unbiased measurements for high-dimensional time-bin-based photonic states. *EPL* **104**, 30003 (2013)
86. Z. Zhang, J. Mower, D. Englund, F. Wong, J. Shapiro, Unconditional security of time-energy entanglement quantum key distribution using dual-basis interferometry. *Phys. Rev. Lett.* **112**, 12 (2014)
87. D. Simon, A. Sergienko, High-capacity quantum key distribution via hyperentangled degrees of freedom. *New J. Phys.* **16**, 063052 (2014)
88. T. Zhong, Photon-efficient quantum cryptography with pulse-position modulation. *New J. Phys.* **16**, 063052 (2014)
89. Y. Noguchi, H. Takesue, Implementation of quantum state tomography for time-bin entangled photon pairs. *Opt. Exp.* **17**, 10976–10989 (2009)
90. K.M. Rosfjord, J.K.W. Yang, E.A. Dauler et al., Nanowire single-photon detector with an integrated optical cavity and anti-reflection coating. *Opt. Exp.* **14**, 527–534 (2006)
91. J.M. Donohue, M. Agnew, J. Lavoie, K.J. Resch, Coherent ultrafast measurement of time-bin encoded photons. *Phys. Rev. Lett.* **111**, 153602 (2013)
92. F. Bussières, Y. Soudagar, G. Berlin, S. Lacroix, N. Godbout, Manipulating time-bin qubits with fiber optics components, in *2006 Digest of the LEOS Summer Topical Meetings* (2006), pp. 22–23
93. A. Schreiber, K.N. Cassemiro, V. Potoček et al., Photons walking the line: a quantum walk with adjustable coin operations. *Phys. Rev. Lett.* **104**, 00502 (2010)
94. M. Szegedy, Quantum speed-up of Markov chain based algorithms, in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science* (2004), pp. 32–41
95. A. Childs, Universal computation by quantum walk. *Phys. Rev. Lett.* **102**, 180501 (2008)
96. A. Peruzzo et al., Quantum walks of correlated photons. *Science* **329**, 1500–1503 (2010)
97. J.O. Owens, M.A. Broome, D.N. Biggerstaff et al., Two-photon quantum walks in an elliptical direct-write waveguide array. *New J. Phys.* **13**, 075003 (2011)
98. A. Schreiber, A.G. abris, P.P. Rohde et al., A 2D quantum walk simulation of two-particle dynamics. *Sci. Mag.* **336**, 55–58 (2012)
99. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, Quantum cryptography. *Rev. Mod. Phys.* **74**, 145–195 (2002)
100. M. Horodecki, M. Piani, On quantum advantage in dense coding. *J. Phys. A* **45**, 105305 (2012)
101. C.H. Bennett, S.J. Wiesner, Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.* **69**, 2881–2884 (1992)
102. T. Das, R. Prabhu, A. Sen(De), U. Sen, Distributed quantum dense coding with two receivers in noisy environments. *Phys. Rev. A* **92**, 052330 (2015)

103. H. Buhrman, R. Cleve, J. Watrous, R. de Wolf, Quantum fingerprinting. *Phys. Rev. Lett.* **87**, 167902 (2001)
104. D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, A. Zeilinger, Experimental quantum teleportation. *Nature* **390**, 575–579 (1997)
105. K. Azuma, K. Tamaki, H.-K. Lo, All-photonic quantum repeaters. *Nat. Commun.* **6**, 6787 (2015)
106. T. Li, F. -G. Deng, Heralded high-efficiency quantum repeater with atomic ensembles assisted by faithful single-photon transmission. *Sci. Rep.* **5**, 15610 (2015)
107. N. Sangouard, C. Simon, H. de Riedmatten, N. Gisin, Quantum repeaters based on atomic ensembles and linear optics. *Rev. Mod. Phys.* **83**, 33–80 (2011)
108. H.-J. Briegel, W. Dür, J.I. Cirac, P. Zoller, Quantum repeaters: the role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932–5935 (1998)
109. M. Zukowski, A. Zeilinger, M.A. Horne, A.K. Ekert, “Event-ready-detectors” bell experiment via entanglement swapping. *Phys. Rev. Lett.* **71**, 4287–4290 (1993)
110. J.-W. Pan, D. Bouwmeester, H. Weinfurter, A. Zeilinger, Experimental entanglement swapping: entangling photons that never interacted. *Phys. Rev. Lett.* **80**, 3891 (1998)
111. E. Megidish, A. Halevy, T. Shacham, T. Dvir, L. Dovrat, H.S. Eisenberg, Entanglement swapping between Photons that have never coexisted. *Phys. Rev. Lett.* **110**, 210403 (2013)
112. A.M. Goebel, C. Wagenknecht, Q. Zhang et al., Multistage entanglement swapping. *Phys. Rev. Lett.* **101**, 080403 (2008)

Chapter 6

The Role of Photon Statistics in Visual Perception



Leonid Krivitsky and Vadim Volkov

Abstract We address the question of how fundamental photon fluctuations are perceived by a live visual system. The discussion is focused on specific type of photoreceptor cells within the eye, known as retinal rod cells. Rod cells provide vision under low light conditions and they are sensitive at a single photon level. We review experiments on interaction of the rod cells with light sources of different photon statistics, including coherent, pseudo-thermal, and single-photon sources. Accurate control over photon statistics of light stimuli, combined with technique for the readout of rod cells response, enable precise and unambiguous characterization of intrinsic features of the visual system at single and discrete photon levels.

6.1 Retinal Rod Cells

The eye represents a unique device for visual perception that developed over millions of years of evolution. Its mode of operation is perfectly tuned and well organized. In fact, some of the features of modern optical engineering can be readily found in live visual systems. Examples include broad-band polarization retardation plates in the eyes of crustacean, polarization sensitive vision of desert ants that helps to navigate under a clear sky, gradient index lenses found in compound eyes of insects that minimize aberrations, and many others.

L. Krivitsky (✉)

Institute of Materials Research and Engineering, Agency for Science Technology and Research (A*STAR), Singapore, Singapore
e-mail: Leonid_KRIVITSKIY@imre.a-star.edu.sg

V. Volkov

Faculty of Life Sciences and Computing, London Metropolitan University, London, UK
e-mail: vadim.s.volkov@gmail.com

V. Volkov

Department of Plant Sciences, University of California, Davis, CA, USA

Specialized cells within the eye, known as retinal photoreceptors, are responsible for sensing light by converting it into the form of electrical pulses. In vertebrates photoreceptors are arranged in a layer, which forms a part of multi-layered retina tissue at the back of the eye. The retina amplifies visual signals and transfers them to neurons of the optic nerve for further processing by the brain. There are two types of photoreceptor cells: cone cells are responsible for colour vision in day light conditions and more sensitive rod cells provide uncoloured vision under low light conditions [1].

The present study is inspired by the ultimate sensitivity of retinal rod cells down to the single photon level. This remarkable property makes them a perfectly suited system for fundamental studies of the role of quantum effects in photochemistry, neurobiology and perception. From the technological standpoint it is highly intriguing that a single rod cell of an average length of about 50 μm and diameter about 5 μm represents a self-contained single photon detector, which includes a light sensitive pigment, an ATP power supply, and a synaptic terminal that links it to the rest of retina. Such a compact arrangement surpasses modern man-made devices which are more bulky and often have technological limitations. Further insights into functions of rod photoreceptors could define the properties required for a new family of sensitive light sensors, mimicking natural detection.

6.1.1 Morphology of Rod Cells

Vertebrate retinal rod photoreceptor cells have a typical rocket like shape, see Fig. 6.1. They consist of two distinctive morphological functional regions: extended rod-like outer segment (ROS) which is filled with photopigment molecules represented by rhodopsin, and shorter rounded inner segment (RIS) which contains components of cell machinery.

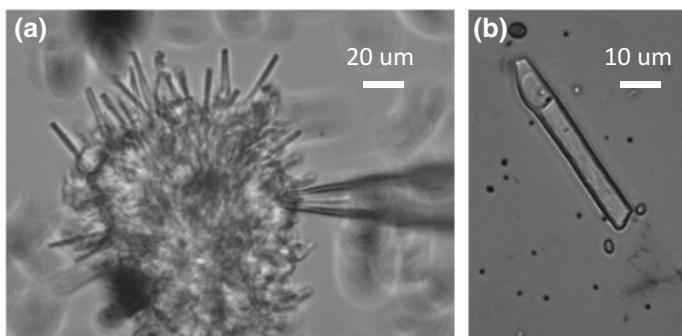


Fig. 6.1 Microscope image of isolated *Xenopus laevis* retinal cells (a) and a rod cell from the preparation (b). Adapted from [105]

Typical size of rod cells is about 2–10 μm in diameter and about 15–100 μm long, the values depend on taxonomic group of animals. Rods of mammals are smaller in diameter than amphibian ones, about 2 μm compared to 5–8 μm , correspondingly. The diameter of rods in retina of birds is also about 2 μm (e.g. for owls [2]), close to the estimated theoretical limit (based on morphometric analysis) for the diameter of rod cells to be 1.2 μm [3]. Still the typical diameters of rods of e.g. African frog *Xenopus* are about 10 wavelengths of red light. Technically speaking, the theoretical resolution of human retina with rods being about 2 μm in diameter is 12,550 dots per inch. The value is comparable to state-of-art charge-coupled device (CCD) cameras, and is sufficient to resolve the tiny details of the surrounding world. Likely, that the high resolution of the retina is naturally restricted by the wavelength of the visible light.

6.1.2 Light Detection by Rod Cells

The basis for light detection by the rod cells is as follows. Rod cells contain large number of pigment molecules in their outer segment. For example, mouse rod contains nearly hundred millions of rhodopsin molecules (reviewed in [4]). A pigment molecule absorbs photon and changes its conformation; the change is amplified by molecular networks of the corresponding rod cell and may result in physiological responses.

More details require knowledge about the structure of rod cells and molecular and physiological mechanisms of signal amplification and transduction in visual receptors. Oblong rod cells have numerous stacked disks at the ROS; the disks are formed by membranes with a rhodopsin photo-pigment [1]. Large number of disks, e.g., about 800 for a mouse rod [4] or about 1500–2000 for an amphibian rod cell [5], multiplied by huge number of rhodopsin molecules per the disk, about $8 * 10^4$ per a disk in the mouse rod, results in nearly hundred millions of rhodopsin molecules per a typical mouse rod [4] or more, about $3 * 10^9$, per an amphibian rod cell [5]. A molecule of rhodopsin consisting of retinal and protein part opsin absorbs photon, changes conformation and isomerizes to metarhodopsin.

Metarhodopsin has a short life half-time, hence special precautions were taken to crystallize this G-protein coupled receptor and solve its structure [6] adding to the earlier crystal structure of rhodopsin [7] and understanding the transformation of rhodopsin after absorption of a photon.

Amplification of a signal from activated rhodopsin occurs via further activation of G-protein transducin, see Fig. 6.2. Transducin is composed of α , β and γ sub-units, it laterally diffuses on the surface of disk membrane and interacts with metarhodopsin. The result of the interaction is that transducin changes bound GDP for GTP (reviewed in [8–10]). Activated α -subunit-GTP of transducin binds phosphodiesterase (PDE) with stoichiometry 2 to 1, thus activating PDE. PDE is

enzyme which hydrolyses cyclic nucleotide cGMP to nucleotide GMP, hence concentration of free cGMP in cytoplasm of a rod cell and more in ROS decreases nearly two times from about 2–4 μM [11–13]. The drop in cGMP closes cyclic nucleotide gated ion channels (CNGCs) of rod cell membrane, because the ion channels are regulated (gated) by bound cGMP [14, 15]. The channels are in an open state in darkness under higher μM concentration of cGMP. Lowered cGMP under illumination closes them in a strong concentration-dependent manner due to four cooperative sites of cGMP binding for the functional channel molecules (reviewed in: [13, 16]). Rod cell membrane hyperpolarizes with closed CNGCs. Finally the initial absorption of a photon results in the decrease of membrane potential to more negative values and gives rise to corresponding ion current, the hyperpolarisation is further passed to neurons of visual nerves (reviewed in: [8–10]).

In reality the basic scheme of physico-chemical events from photon absorption by rhodopsin to membrane depolarisation in rod cell due to closure of CNGCs is much more complicated. It includes numerous feedbacks, mechanisms of regulation, signal/noise suppression and stable and robust amplification of signal. Without excessive over complication we can calculate basic parameters corresponding to absorption of one photon, see Fig. 6.3. The quantum yield of rhodopsin transformation was estimated over 0.6 (reviewed in [17]), then one molecule of metarhodopsin R* can activate up to hundreds of transducin (G* for active form) molecules [8, 18]. Rate of activation is around $125 \text{ G}^* \text{ s}^{-1}$ per R* for amphibian rods at room temperature and about 3 times higher in mammalian rods at body temperature [10]. Transducin activates PDE (ratio 1:500 was proposed for R* to activated PDE [18]) and finally up to 10^5 cGMP molecules are hydrolyzed per photolyzed rhodopsin [5]. Further estimates include volume V for a typical retinal rod cell (cylinder with diameter about 5 μm and length about 50 μm) of *Xenopus* toad being about $1000 \mu\text{m}^3 = 1 \text{ pL} = 10^{-15} \text{ m}^3$ and surface area S around $1600 \mu\text{m}^2$. Hence, about 10^5 cGMP molecules are equal to concentration of about $10^5 / (V * N_a) = 10^{17} / (6.02 * 10^{23}) \approx 0.16 \mu\text{M}$, where N_a is Avogadro constant. The real changes in the concentration of cGMP after absorption of a single photon are to be higher and located within a much smaller volume of the ROS.

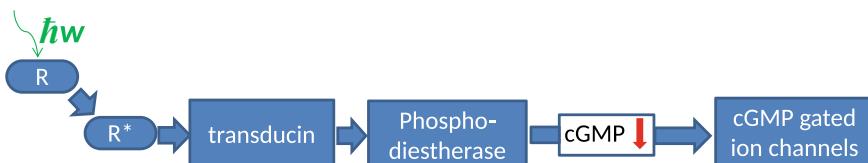


Fig. 6.2 Signal transduction chain in a vertebrate retinal rod cell starting from a photon $h\nu$ and leading to cyclic nucleotide gated channels. Closure of the ion channels after drop in cGMP results in membrane hyperpolarisation and stops inward ion current of sodium and calcium. Adapted from [105]

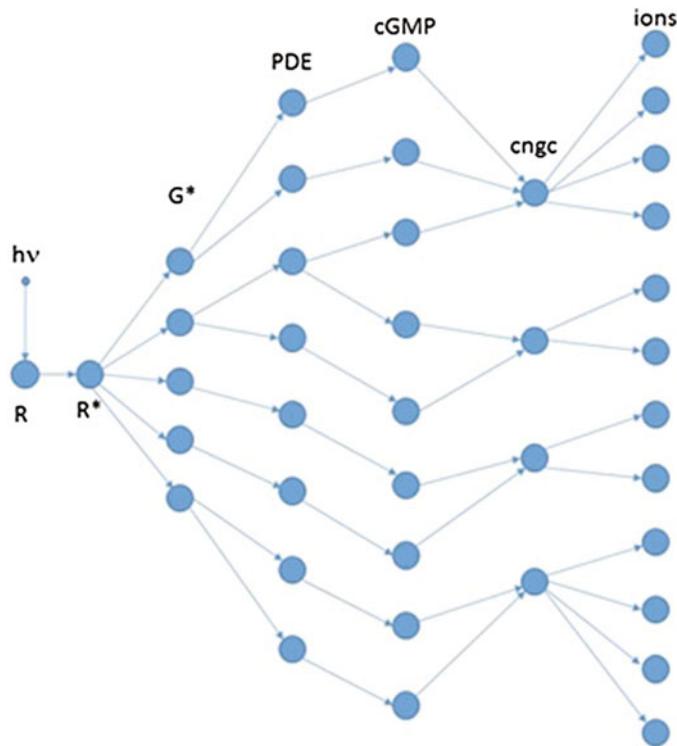


Fig. 6.3 Directed graph of interacting elements for modelling signal transduction in a vertebrate rod cell after influence of one absorbed photon. R is rhodopsin, R^* is activated rhodopsin, which activates over 100 transducin G^* molecules and in turn about 500 phosphodiesterase PDE molecules. PDE reduces concentration of cGMP for about 10^5 molecules/ R^* or about twice from 2 μM , drop in cGMP closes cGMP-gated ion channels (CNGCs); it changes membrane potential and ion fluxes via membrane. The number of affected molecules is indicative and was determined in multiphoton experiments, where changes depend on duration and intensity of light stimulus. Not all the components are included for initial simplicity (e.g., Ca^{2+} signalling, lipid signalling etc.), while further interactions in prolonged multiphoton experiments may include alteration in gene expression and physiological state of the rod cells. Not all the interactions are confirmed for single photon experiments

6.1.3 Mechanisms of Reproducibility for Light-Induced Responses of Rod Cells

The molecular machinery for detection of photons by rod cells (1) should provide that the electric response is proportional to the number of absorbed photons, (2) has to ensure feedbacks to quench the activated rhodopsin (metarhodopsin R^*) and (3) needs to terminate the signal amplification. Without these mechanisms a single photon would trigger a chain reaction in the cell and the light signal would be distorted, thus providing false information about the real world.

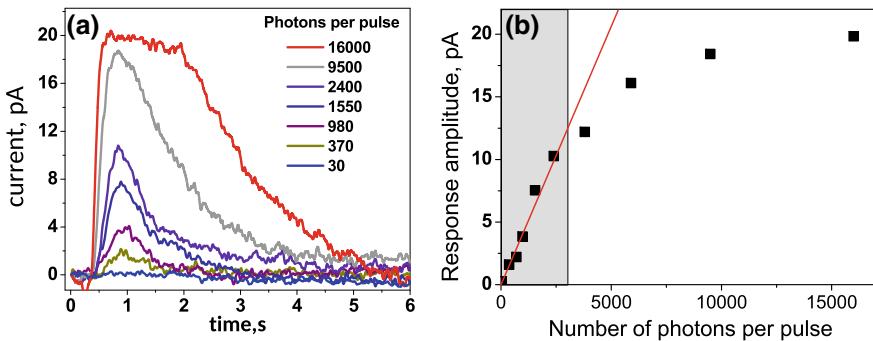


Fig. 6.4 **a** Kinetics of retinal rod responses of *Xenopus laevis* to pulses of light. The colour legend indicates the corresponding number of impinging photons, which were emitted from the tip of the optical fiber (number of photons is not adjusted for the efficiency of light delivery to the cell). **b** The corresponding dependence of the amplitude of the cell response on the number of impinging photons. The grey shaded area shows the region of the linear response. From Sim et al. [99]. Results for this figure and for Figs. 6.5, 6.6, 6.7, 6.8, 6.9, 6.10, 6.11 and 6.12 are for the rod cells from dark adapted retinas of frogs *Xenopus laevis*

Increase of the number of photons impinging on a rod cell increases the electric current recorded from the cell in a reproducible manner, Fig. 6.4a, b. The higher number of photons activates more rhodopsin molecules, hence the response amplitude is higher; the response amplitude reached saturation at around 25–30 absorbed photons per light pulse for *Xenopus* rod cells [19].

Molecular networks for termination of signal in rods include many proteins and protein macromolecular complexes. Under dark conditions metarhodopsin, activated transducin and PDE are deactivated, then concentration of cGMP returns to earlier higher values resulting in depolarization of the membrane. These processes determine temporal and kinetic components of the light-induced events in rods (reviewed in [20]). Inactivation involves several fast steps: for example, in mouse rods the activity of R^* is quenched with half-time about 50–80 ms [21, 22] by successive phosphorylation by rhodopsin kinase and further binding of the protein arrestin [20]. Each of the proteins, rhodopsin kinase and arrestin, in turn has their own regulation feedbacks and loops.

The kinetics of light-induced photocurrent in rods, shown in Fig. 6.4, can be reasonably well described by several mathematical models, which account for kinetics of individual photon-induced reactions, kinetics of inactivation, and the morphology of rod cells [20, 23, 24]. It is worth to mention that reproducibility of electric output from a rod cell after absorption of a single photon is remarkably stable. At a first glance it seems that the numerous stochastic probabilistic components would make the response also stochastic with high variability. However, the numerous feedbacks, cooperativity of interactions and buffering of concentrations by diffusion rates make the responses relatively robust [20]. Inactivation components are especially important for ensuring the reproducibility [24, 25]. Modelling was performed using detailed geometry to imitate rod disks as circles

with known incisions [24]. Results led to conclusions that diffusion rates of cGMP and the other second messenger and regulator Ca^{2+} in cytoplasm are the main suppressors of variability which occurs in the activation cascades [24]. Concentration of Ca^{2+} ions is one more regulator of cGMP-gated ion channels (reviewed in: [26]). Charged amino acid glutamate residue is located in the pore of the protein which forms the channels; this amino acid residue is responsible for block of the inward current through ion channels by micromolar Ca^{2+} concentrations [27]. Ca^{2+} ions are also passing via the channels and add an additional feedback regulation loop.

An important role in optimization of the signal-to-noise ratio in rod cells is played by cooperativity of interactions and by low electric conductance of cGMP-gated ion channels. The natural choice was to select a large number of ion channels with abnormally low conductance, hence thermal fluctuations of gating for a single ion channel will be averaged and have no significant effect on the membrane potential. About 500 cGMP-gated ion channels are estimated per μm^2 in salamander rods making the total number around $5 * 10^5$ per the cell with a small fraction of 1–2% being open even in darkness (reviewed in: [28]). A single channel has tiny conductance of 100 fS or approximately 3 fA at -30 mV in a standard Ringer medium [14]; (reviewed in: [28]). Several hundreds and thousands of channels are required for response to one photon and multiphoton pulse, respectively. It allows the passage of about $2 * 10^4$ monovalent cations per second per a single channel at -30 mV. One more, but slower component of noise originates from transduction cascade due to spontaneous thermal activation of PDE and the other component originates from thermal activation of rhodopsin with low probability of 10^{-10} (reviewed in: [29]).

Our present knowledge and recent progress in molecular biology already offer opportunities to manipulate the light-induced electric responses of a rod cell. There are numerous mutations influencing components of the signal transduction chain; some mutants have altered kinetics and amplitude of photocurrent (reviewed in [20]). For example, C-terminus of rhodopsin molecule has six sites for phosphorylation, they are important for inactivation of R^* . Decreasing the number of the sites in mouse mutants by means of molecular biology increased the duration of photocurrent and also changed its shape [30]. It is interesting to mention that phosphorylation at amino acid residues of serine or threonine had distinct effects on photocurrent curves [31]. Earlier observations on the role of Ca^{2+} ions in photo-transduction networks led to the opportunity of shaping the photocurrent by buffering Ca^{2+} in cytoplasm of rod cells. Ca^{2+} -chelator BAPTA kept stable Ca^{2+} concentration; it slowed $\text{Na}^+/\text{Ca}^{2+}$ exchange current via transporters of rod cell and hence increased amplitude and duration of the total photon-induced photocurrent [32]. Less directed option of using specific inhibitors or stimulators of cell biochemistry was realised in the experiments at the beginning of 1980s. Surprisingly, inhibitors of PDE increased the electric photoresponse of rod cells though the opposite is expected according to our present knowledge [33]. The voltage change of several mV was recorded by intracellular electrodes in rod cells after light pulses. PDE inhibitors including caffeine, papaverine and several others enhanced two to

six times the voltage photoresponses [33]. Potentially these effects are linked to Ca^{2+} concentrations and need analysis from the point of systems biology. Even simpler option of changing extracellular K^+ or Cl^- concentrations has an effect on voltage decrease in rods after pulses of light [34].

6.1.4 Comparison of Photoreceptors from Different Organisms and Man-Made Photodetectors

One important note is that different biological organisms possess slightly different phototransduction chains and often quite large variation in parameters of light-induced electric responses. For example, human rod cells operate at stable body temperature, smaller in size than amphibian rods, but have shorter time to peak of the photoresponse with about 100-fold higher calculated amplification coefficient (reviewed with more comparisons in: [35]). The operating temperature and morphology could be an explanation for the phenomenon [36]. Comparative studies of retinal rods of many vertebrates revealed significant variation in amplitude of single-photon responses from 0.4 pA with signal/noise ratio (SNR) around 1.5 (amplitude over standard deviation of the biological component of dark noise) in river lamprey to 0.4–1.1 pA with SNR 1.6–3.7 in mouse according to different experimental data (reviewed in [37]). The higher SNR about 4.5 for the photoresponses was reported for rods of monkey *Macaca fascicularis* due to low background noise of rods [36].

More comparisons lead to invertebrates. Insects have no retina and rod cells, but compound eyes with light perceiving cells called ommatidia. Surprisingly, insects have similar to vertebrates phototransduction chains and architecture of retinal neurons [38], though with slightly different sequence of events for phototransduction. Fruitfly *Drosophila* is a well-known biological insect model with numerous mutants available for understanding the light perception in the ommatidia of the organism. The known differences between a fruitfly and vertebrates in phototransduction is that (1) phospholipase C is present in photoreceptors of *Drosophila* instead of PDE in vertebrates, (2) signalling via inositol trisphosphate and diacylglycerol and probably polyunsaturated fatty acids in fruitfly substitutes cGMP signalling in vertebrate rod cells, (3) phototransduction in photoreceptors of the insect results in opening of closed under darkness transient receptor potential ion channels, not closing of cGMP-gated ion channels in vertebrates (reviewed in: [39–41]). Kinetics of photoresponses to single photons in fruitfly is 10–100 faster than in vertebrate rods [42]. More variation may be found among species of numerous and strikingly unusual biological organisms.

The relatively low SNRs for vertebrate rod cells seem to be typical for biological systems compared to higher SNRs of man-made photo-detection devices. It's reasonable to analyse the sources of noise in rod cells and in engineered photodetectors. The main component of background high frequency noise in rod cells is

determined by flickering (opening-closing) behaviour of open cGMP-gated ion channels. Suggesting normal or binomial distribution for flickering amplitudes of the large number of ion channels [43, 44], it is conceivable that the noise is proportional to electric conductance of a single channel and the number of the channels. From one point, general assumptions propose limits for noise suppression in biological systems, otherwise negative feedbacks and the whole functioning of system are becoming extremely expensive: the minimum standard deviation decreases with the quartic root of the number of events for Poisson communication channels [45]. From the other point, noise could be important for the cell behaviour [46] and for interaction with neurons. Hence, it is likely that the background flickering noise of ion channels could not be essentially reduced. The single-photon signal at the background of the noise is resolved by the amplification cascades after activation of the rhodopsin, so the SNR is the result of two independent processes. Considering for instance a ten times lower SNR would prevent from recognising single photons by rod cells; ten times higher SNR would essentially shrink the range of multiphoton response since the maximal photocurrent is limited by interactions with neurons. Again, the design of rod photo-responses seems optimal. The slower sources of noise are generated by transduction cascade due to spontaneous thermal activation of PDE or thermal activation of rhodopsin with low probability of 10^{-10} [29, 47], the latter cannot be distinguished from single photon responses.

Man-made photodetectors differ in the photosensitive elements, from (1) alkali and group V metals or their alloys in photomultiplier tubes, (2) silicon or germanium photosensitive semiconductors in avalanche photodiode detectors to (3) silicon-based photodetectors in charge-coupled devices (CCDs) [48] and to (4) Niobium nitride nanowires for superconducting photodetectors [49]. Consequently, the quantum yield varies from 0.2 to 0.95 [48] compared to about 0.7 of rhodopsin. SNRs of man-made photodetectors at the single-photon level also differ and could be below 1 in Geiger-mode avalanche photodiodes [50, 51], while reaching over 100 at high light intensities [50]. More details are briefly summarised in Table 1 and potentially may provide ideas to improve the man made photodetectors.

6.1.5 *Basics of Communication Between Neurons and Rod Cells*

Change of electric potential in a rod cell is sensed by neurons in the retina and directed for further processing to specific regions of the brain. The giant “neural supercomputer” consists of: hundreds of neurons in worms lacking eyes, thousands and millions of neurons in insects with compound eyes and over 80 billions of neurons in the human brain [52–55]. The processing of visual information provided by photoreceptors (including retinal rod cells) is extremely complex and essentially

Table 1 Comparison of basic photodetection properties of rod cells with man-made engineered photodetectors

	Rod cells	Photomultiplier tubes	Avalanche photodiodes
Photosensor compound	Rhodopsin	Alkali and group V metals	Silicon, germanium-based or other semiconductors
Quantum yield/efficiency	About 0.7	Usually 0.1–0.5, over 0.7 for GaAsP	Usually 0.2–0.4, though could be over 0.8 for silicon-based and other semiconductors
Transformation of signal	Photon changes the conformation of protein molecule	Light signal is directly converted to electric signal	Light signal directly excites the electron-hole pair
Number and properties of amplification cascades	Several steps including at least four amplification cascades with different properties: amplification of protein conformation, of hydrolysed cGMP, of cation current	Several multiplication cascades to amplify the initial electric current of electrons	Direct amplification of initial electric signal through impact ionization
Principle of interactions in amplification cascades	Several mechanisms including: (1) molecular interactions of diffusing proteins, which precisely recognise the corresponding proteins from the next amplification cascade; (2) chemical reactions; (3) binding of cGMP to protein ion channels; (4) final step is conversion to electric signal of ion fluxes	Electrons emitted from photocathode are multiplied by several dynodes via secondary emission of electrons	Impact ionization by electrons/holes in the multiplication region under strong electric field
Response time	Tens and hundreds of milliseconds to few seconds, high variability among species and Q_{10} in physiological range of temperatures	Around 1–20 ns depending on type, recovery time is below and around 100 ns	Around and below 1 ns, recovery time is below 50 ns

Based on Ref. [48] and manuals for present photomultiplier tubes and avalanche photodiodes

not known yet. Changes in electric potentials of individual rod cells are passed to neurons of adjacent layer in retina, then to the second layer of specialised neurons and finally to neurons of visual nerves (more details are reviewed in [56]). After several consecutive steps the electric pattern of altered electric activity is processed, modified and analysed by brain. The processing and modification involve correction for image aberrations, uneven distribution of photoreceptors in retina and the other drawbacks of initial signals. However, the question rises if the electric signals from individual rod cells are reliably sensed by neurons and also provide required starting level of temporal and spatial resolution. Indeed, a photoresponse from a rod after absorption of a single photon could be too low or too high for retinal neurons. To the best of current knowledge natural retinal design is essentially the mostly optimal [57].

While the photoelectric response from a rod could be spread by means of electrically coupling the rod cells via special gap junctions of high resistance (reviewed in [56]), the same signal could be amplified in synaptic terminals of rod cells. The electric coupling between rods in the retina reduces temporal resolution, but increases signal to noise ratio; the gap junctions are the regulated contacts and potentially they offer a way to adapt to changing illumination.

Synapses are the other specific places of contacts between cells; synapses of nervous system are divided to chemical and electric ones. Briefly saying, chemical synapses are composed of two membranes of contacting cells and have a small gap of around 20 nm between the membranes. The specific chemical compounds are released by one membrane, pass the gap within milliseconds and interact with receptors or specialised ion channels at the other membrane known as postsynaptic one. It leads to depolarisation or hyperpolarisation of the corresponding neuron, which forms the postsynaptic membrane. The simple mechanisms links cell biology with electric responses of neurons, moreover it allows amplification of the signals. There could be many synapses between two adjacent neurons. Finally a “giant supercomputer” is formed consisting of numerous neurons with several potential discrete or gradual states of electric potential. Changes in electric potential are able to carry, record and process information. Recent progress in computer modelling helped to simulate the behaviour of about 31,000 real neurons of 207 revealed subtypes with approximately 8 million connections and 37 million synapses [58].

The change in electric potential of a rod cell after a response to single photon is reported from about 0.2 mV/R* to over 1.0 mV/R* [34] or similar values of 1 mV per a single absorbed photon [59]. The voltage changes correspond to measured ion currents about 0.3–1.5 pA/photon (reviewed in: [25, 37]). Small fluctuations of rod signals are effectively filtered in retinal bipolar cells owing to strong nonlinearity in synapses caused by specific ion channels [57, 60]. About 0.5 to 2 bipolar channels in transduction chains of synapses are opened in darkness whereas about 30 channels are open after response to one photon [60]. Thus the optimal natural design efficiently suppresses the noise from rods [57]. One more interesting feature is that too strong signals above a certain threshold from rods are clipped by voltage gated calcium channels in synapses. It prevents from excessive voltages and, hence, the large dynamic range of photon fluxes is compressed within a narrow range of

voltages of about 5 mV [61]. However, ion currents in synapses are highly dependent on the external pH [62] and the synapses are very dynamic structures with numerous regulators [63]. It was proposed that the features convert the photovoltage range from rods for nonlinear processing by neural networks [64].

6.1.6 Readout of Rod Cell Response

Historically, response of photoreceptors to weak light flashes has been intensively studied since the 1930s with first experiments dating back to 1889 [65] and 1907 [66]. The main motivation of those studies was to determine the fundamental limit of light perception in nature and to understand its variation in different species. The experiments coincided with the prediction and discovery of photons at the beginning of the 20th century, quantum physics gave explanations and independently stimulated study of visual receptors. Early physiological approaches were based on so-called “frequency-of-seeing” experiments, when a dark adapted individual was asked about the perception (“can see”/“can not see”) of faint light flashes. The probability of seeing a flash was recorded for light pulses of different intensity. From fitting the observed probability curve it was possible to deduce the visual threshold, which turned out to be in the range of 2–8 photons impinging on the retina. The observations are back to ancient astronomy where we know the flux of photons from distant stars and able to correlate the visibility of a star with the corresponding photon flux.

More precise suitable methods appeared later. Intracellular recordings in electrophysiology appeared by 1950s after introducing glass microelectrodes with tiny thin sharp tips (diameter of about 0.1 μm) (history of electrophysiology is well reviewed in: [67]). These impalement electrodes gave opportunity to insert the sharp tips into a single cell and measure electric properties of cell membranes under certain conditions. This method is very productive to measure kinetics of membrane potential in rod cells under illumination. After the initial recordings with the preparations of retina (e.g. [68]) it was suggested that the main source of electric activity in retina under illumination were rods and cones [69]. The study shifted to individual rod cells [70, 71]; from the other side the use of arrays of extracellular electrodes confirmed that the photoelectric activity of retina is linked to outer segments of rod cells [72]. Application of numerous inhibitors and stimulators together with varying the external medium for rod cells provided plenty of information about the mechanisms of the processes [33, 34]. By the end of the 1970s, with an invention of the patch clamp technique more opportunities become available. Indeed, the technique allows to change the intracellular composition and record ion currents under determined voltages. It helped to decipher the role of cGMP in regulating ion currents [14] and completed the scheme of phototransduction chains and intracellular events.

The other approach was also developed allowing direct measurement of membrane current of a single photoreceptor upon light stimulation [73, 74]. The method,

known as suction electrode, uses tight glass micropipette to monitor the membrane current of the rod cell. One segment of the rod cell is drawn in the micropipette filled with the physiological solution, see Fig. 6.10b. The rod cell is functioning as *in vivo*, but the current, flowing through the cell membrane, is now re-directed to a low noise ampere meter. In the dark the microelectrode measures the current, caused by continuous transport of ions (Na^+ , K^+ , Ca^{2+}) through ion channels in the membrane. Once a photon hits the ROS, it isomerizes the rhodopsin pigment and starts a phototransduction cascade, which results in a closure of specific ion channels and prevents ions from entering the cell. Change of the membrane current (typically a few pA) is detected by the suction micropipette. The examples of waveforms of the rod photocurrent at various number of impinging photons are shown in Fig. 6.4.

Methods of delivering regulated number of photons to eye or rod cells also progressed since the end of the 19th century: the parallel or biased advances in (1) optical methods and theory from one side and (2) ways of recording response from rod cells from the other side determined the trajectory of the research.

6.2 Overview of the Earlier Visual Experiments

The first known reported visual experiments are dated back to 1889 [65]. The aim of the research was to determine the sensitivity of eye depending on the colour of light. The light energy was measured by self-made sensitive bolometer, while an observer indicated the visibility of attenuated light ray. Four different observers were chosen for experiments. The conclusions were that sensitivity to green light was about 100 times higher than for red or violet light and the green light sensitivity corresponded to about $3 * 10^{-9}$ erg or $3 * 10^{-16}$ J. Recalculating we assume that about 1,000 photons were the minimal visible number for the experiments, a reasonable initial result taking into account numerous sources of errors described in the paper [65]. Further experiments led to about 35–70 photons [66] or about 40–90 photons as a minimal sensitivity threshold in these experiments [75] (several experimental results are summarised and reviewed in [76]).

It is worth mentioning and describing in more detail the experimental results of Sergey Vavilov (Wawilow) on visual perception obtained in 1930s and 1940s since the results are rather undervalued yet [77–81] (Translation into English of some of his papers is included in the Part 2 of this book). The thoughtful and well-organised experiments included dark preadaptation of eye for about one hour and used uranium glass and several uranium salts to check the level of adaptation. Dark adapted eye was able to see the luminescence of uranium glass. Special fixation point was helpful to set the defined position of eye and direct photons to specific most sensitive part of retina. The excitation of the eye lasted for 0.1 s with further gap for 0.9 s. The observer did not know the presence or absence of excitation by photons. The minimal number of detected photons was not a strict number with sharp threshold, but varied depending on the observer and even changed twice within the

same day for the same observer following physiological conditions. The number was considered as photons impinging on the cornea of eye in the first experiments of 1930s and later on in 1940s as photons absorbed by retina. Plotting the probability of seeing against number of photons helped to determine the visual threshold for retina based on the slope of the curves. The threshold was about 20–40 photons (from 8 to over 45 for different observers) for several series of experiments over 10 years. The source of light was an incandescent light bulb supplied by electric power, the light beam was weakened in a quantitative manner after the green filter to ensure the highest eye sensitivity. Since eye was the most sensitive light detector at the time, eye was also used to detect quantum fluctuations in weak light beams, hence served as a measuring device for precise physical measurements [77–81].

Among classic behavioral experiments on threshold of vision are the works of Hecht and colleagues [76, 82]. The researchers determined the minimal number of detected photons of green-blue range of spectrum as 54–148 and made assumption concerning the number of photons reaching retinal rods. Estimates for absorption and reflection by cornea, absorption by lens and the other parts of eye to around 90% of the initial value allowed to find the minimal threshold of photons sensed by retina: the threshold was from 5 to 15 photons.

Behavioral experiments are very important for understanding general peculiarities of vision and also for setting background for further more detailed visual research. To compare with human visual threshold the behavioral study included even animals, e.g. cats and owls; cat had 6 times lower visual threshold [83], owl also likely had a lower threshold compared to human, while having similar spectral sensitivity [84].

Interesting and highly analytical series of behavioral experiments was carried out with laser source of light [85–87]. Plane polarised light pulse with duration 1 ms from Ar⁺ green laser had Poisson distribution of photons. The 1 ms pulse was used for “can see-can not see” trials with four trained male observers. Several options were offered for the observers from (1) reporting positive and negative responses of seeing the light pulse (low false-positive rate) to (2) encouraging them to report on any possible occasion of seeing a light flash (high false-positive rate). Low false-positive rate corresponded to 127–147 photons as minimal number of visible photons impinging on cornea, high false-positive rate gave threshold being 42–82 photons. The obvious conclusion is that reliability and sensitivity are not compatible. Modelling gave numeric parameters: 60% probability of seeing with 1% of false-positive rate gave 147 photons, rise of false-positive rate to 33% decreased the number of photons to 34 [85]. Assessing the experimental data resulted in the threshold at retina from 11 to 32 photons with similar values of dark noise at retina [86]. Moreover, the next important step was to evaluate the role of photon statistics from the light source on the visual perception [87]. Intensity modulation of a Poissonian laser light gave nearly flat distribution with similar number of counts per unit of time instead of peak in Poissonian distribution; it changed the curve for probability of vision for the same energies. At low false-positive rate the corneal threshold increased from 147 to 162; the curves at both high and low false-positive

rates became more shallow [87]. The phenomenon was explained by higher frequency of trials with low number of photons and higher uncertainty compared to Poissonian distribution. The researchers discussed the ideal unibunched light with sub-Poissonian statistics for potentially getting more information for probability curve and threshold of vision. From one side, these detailed experiments provided further understanding of visual perception depending on (1) noise with different distributions of events in visual system, (2) statistics of photons and (3) false-positive settings [86–88]. From the other side, they obviously pointed to the limitation of behavioral experiments due to high losses in eye and dependence on physiological conditions of observers.

The mainstream of research evidently and inevitably led to complementing study of isolated rod cells [14, 73, 74, 88–90]. Microelectrode recordings (e.g. [91, 92]) and later both patch clamp [14, 30, 31, 89, 90] and suction electrode [73, 74, 88] techniques were productive for the visual experiments. We will describe the experiments briefly since the further part provides similar data using more controlled source of light. Microelectrode recordings demonstrated hyperpolarisation of rod cells from below 1 mV up to 30 mV after illumination depending on the intensity of the light stimulus [e.g. 91, 93]. Similar dose-response curves with saturation for ion currents were recorded by suction electrode for rods of toad *Bufo marinus*. Saturation values for sigmoid curve (logarithmic scale of abscissa for number of photons) of current were around 20 pA, the source of light used was tungsten-iodide lamp [73, 74, 88]. The electric response to dim light was quantised suggesting that an individual rod cell was able to absorb single photons and respond by nearly 1 pA current pulses with quantum efficiency equal to 0.5 ± 0.1 [74]. The experiments also confirmed the spectral sensitivity of rod cells peaking in green light with wavelength about 500 nm [73].

Patch clamp recordings with isolated rod cells added more information about (1) the reversal potential of light-sensitive ion currents indicating to specific cations for the current, (2) properties of light-sensitive noise at high frequencies of recordings and (3) were helpful to measure single photon responses of genetically modified mice rod cells [e.g. 14, 30, 31, 89, 90]. Noisy recordings under dark conditions corresponded to fluctuations in conductance of cyclic nucleotide gated ion channels, light pulse closed these ion channels and noise at the corresponding frequency dropped [90]. Several other components of noise were analysed; it is worth to mention again the large spontaneous fluctuations of dark noise in vertebrate photoreceptors derived from spontaneous isomerisation of rhodopsin, the peaks of current and voltage are similar to single photon responses (reviewed in: [94]). Interestingly, the light-dependent saturation current in frog rods measured by patch clamp was about 25 pA, which is similar to recordings in toad rods by suction electrode [90]. Basic parameters of photocurrents in rods of jawless lampreys also coincided for measurements by suction electrode and using patch clamp, though patch clamp provided more detailed parameters of ion currents [31]. The light sources for the experiments were either tungsten lamps or light-emitting diodes (LEDs) of green spectral range, while statistics of photons was not controlled and was not specially measured.

6.3 Interaction of Rod Cells with Classical and Quantum Light Sources

In nature we often deal with flickering lights. Representative examples are observation of blinking stars and firecrackers. However, even if the light appears stable to the eye it still exhibits a variation in the number of photons, known as photon noise. The number of photons at each moment is not strictly constant but it is rather described by a specific statistical function, which depends on the light source. For example, the photon number distribution for a laser obeys Poissonian statistics. For a thermal source (lamp, star, LED) the photon number distribution obeys Bose-Einstein statistics. Examples of two distributions, for the same average photon number are shown in Fig. 6.5. As expected, the distribution for the laser is much narrower than for the thermal source. The photon noise (or photon fluctuations) is caused by the quantized nature of light and it cannot be eliminated by technical means.

Historically, visual experiments were mainly focused on studies of averaged visual responses to repeatable light flashes. Direct impact of photon noise on the vision process was often unnoticed. However, as we shall see later in this Chapter, careful account for statistical properties of the impinging light is crucial for accurate interpretation of visual experiments. Light sources with controllable statistical properties can be engineered using tools and approaches of modern quantum optics.

In experiments conducted at a single photon level impact of photon fluctuations becomes increasingly important. Indeed, in this case the average number of photons becomes comparable or even less than the corresponding standard deviation. The statistical distribution of responses of rod cells was shown to follow the discrete distribution of photons, which confirmed the single photon sensitivity of rod cells. Recently rod cells were interfaced with a specialised light source which produces

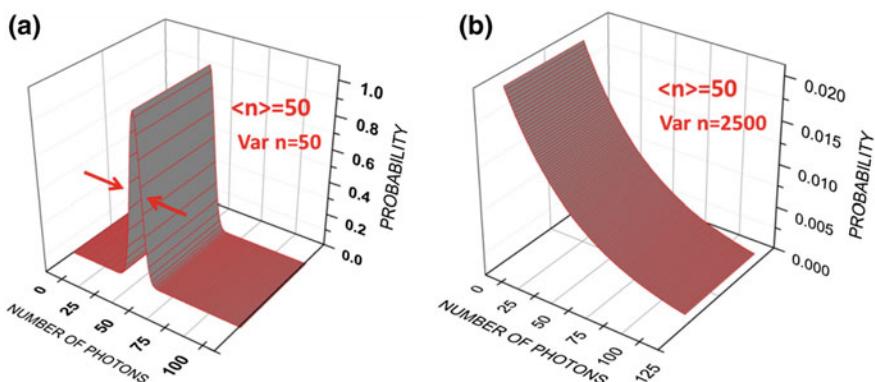


Fig. 6.5 Photon number probability distribution for a coherent (a) and a thermal (b) light sources. The distributions are plotted at the same value of the average photon number of 50 photons. The variance of the number of photons is much narrower for the coherent state of light

either zero or one photon, but never two photons or more [95]. This provided a direct proof of the single photon sensitivity of rod cells.

6.3.1 *Experiments with the Whole Visual System*

Baylor et al. studied responses of rod cells to extremely faint optical pulses, causing the cells to absorb only few photons [73, 74]. Experiments revealed remarkable trial-to-trial stability and discretization of cell responses. These findings led to the conclusion on the ability of individual rod cells to detect single photons. The probability distribution of response amplitudes was found to be Poissonian. However, detailed analysis of influence of photon fluctuations on the cell response was out of the scope of those works.

Influence of controllable photon fluctuations on the response of the visual system was first studied in experiments by Teich and colleagues [85–87]. A light source with a super-Poissonian photon probability distribution was used in conventional frequency-of-seeing experiments. The source consisted of a laser and an acousto-optical intensity modulator (AOM). Setting the shape of the driving signal of the AOM allowed precise control over the shape of the photon number distribution. Such a source mimics the cathodoluminescence light, which can be observed as emission from a fluorophore screen bombarded by electrons.

Frequency-of-seeing curves obtained with the super-Poissonian light source were found to be less steep compared to the curves obtained with a Poissonian source. This effect was associated with difference in photon number distributions. Let us consider the two light sources emit the same average number of photons. The photon number distribution for a super-Poissonian source is broader than for the Poissonian one. Hence the probability of emission of pulses with extremely small and extremely large number of photons is higher for the super-Poissonian source. Pulses with low number of photons are unlikely to be detected by the observer. At the same time pulses with high number of photons are detected with almost the same probability as pulses from a Poissonian source. As a result, the overall probability to see the flash is decreased for the light source with the increased variance in the number of the impinging photons.

The obtained results are particularly important in measurements of the visual threshold, which is conventionally obtained from fitting of frequency of seeing curves. The curves obtained in experiments with light sources of different photon statistics will yield different values of the threshold. Hence, a proper account for statistical properties of the light source is absolutely necessary to avoid any inconsistencies in measurements of the threshold.

6.3.2 Experiments with Isolated Rod Cells and Classical Light Sources

Experiments by Teich and colleagues highlighted the importance of careful attention to statistical properties of light sources used in visual studies. However, the role of individual photoreceptors in the observed effect was hindered because a few hundred of rods were simultaneously illuminated and their collective response underwent several intermediate stages of visual processing. In later experiments Sim and colleagues, individual rod cells were stimulated by light sources with various photon statistics [19].

In their experiments Sim and colleagues, used two examples of classical light sources with noticeably different statistical properties, namely coherent and thermal sources. The photon number distribution $P_{\text{ph}}(m)$ for coherent and thermal sources is given by: $P_{\text{coh}}(m) = \frac{\langle m \rangle^m e^{-\langle m \rangle}}{m!}$, $P_{\text{th}}(m) = \frac{\langle m \rangle^m}{(1 + \langle m \rangle)^{m+1}}$, respectively, where $\langle m \rangle$ is an average number of photons. The statistics of the rod cell response can be derived using a photon counting model. In the model it was assumed, that (1) once the rhodopsin is isomerized, the response occurs with almost unity probability, (2) responses to individual photoisomerizations are additive and (3) responses have a standard Gaussian shape. The theoretical analysis follows three steps:

1. For a given photon distribution $P_{\text{ph}}(m)$, the probability of excitation of n rhodopsin molecules is given by Mandel's formula:

$$P_I(n) = \sum_{m=1}^{\infty} \binom{m}{n} \eta^n (1-\eta)^{m-n} P_{\text{ph}}(m),$$

where η is an overall quantum efficiency of photodetection.

2. Probability for observing a cell response with an amplitude A is given by the average over responses to individual isomerizations:

$$P(A) = \sum_{n=0}^{\infty} \frac{1}{\sqrt{2\pi n \sigma^2}} \exp\left(-\frac{(A - nA_0)^2}{2n\sigma^2}\right) P_I(n),$$

where A_0 is the average amplitude of the response to single isomerization, and σ is its standard deviation.

3. To account for saturation of the response, caused by closure of all the ion channels in the cell membrane, $P(A)$ is truncated at the saturation amplitude

$$A_s \cdot P_S(A) = \begin{cases} P(A), & \text{if } A < A_s \\ 1 - \int_0^{A_s} P(A) dA & \text{if } A = A_s \\ 0 & \text{if } A > A_s \end{cases}$$

From the obtained $P_S(A)$ the k -th statistical moment of the response amplitude can be found: according to definition $\langle A^k \rangle = \int A^k P_S(A) dA$. The signal to noise ratio (SNR) is defined as $\text{SNR} = \sqrt{\langle A \rangle^2 / \text{Var} A}$, where variance is $\text{Var} A = \langle A^2 \rangle - \langle A \rangle^2$. If the amplitude of each response is measured along with a number of photons in each pulse K , the Glauber's second order correlation function $g^{(2)}$, given by $g^{(2)} = \langle AK \rangle / \langle A \rangle \langle K \rangle$, can be obtained [96].

In the experiment two light sources (coherent and thermal) were realized using the same laser (a frequency doubled Nd:YAG laser, modulated by a mechanical shutter) with the wavelength of 532 nm. In the experiment with the coherent source, the laser beam was split on a 50/50 beamsplitter and then directly coupled into two optical fibers, one of which was used for stimulation of the cell and another one was directed to a single photon avalanche photodiode (APD). In the experiment with the thermal source, the same laser beam was first scattered by a rotating ground glass disk, and then, a single scattered speckle was coupled into the fiber. The intensity of the speckle at different orientations of the disc follows the thermal distribution [97].

Dependence of the average amplitude of cell response on the number of impinging photons for the two sources is shown in Fig. 6.6. The curve is much steeper for the coherent source than for the thermal one. This effect can be understood from the analysis of experimental probability histograms of response amplitudes, shown in Fig. 6.7. For stimulation with relatively weak pulses the rod cell behaves as a linear photodetector and the histograms reflect the statistics of light pulses, see Fig. 6.7a, b (red bars). However, once the number of impinging photons is sufficient to cause saturation of the cell response, significant variations are observed. For the coherent source the photon probability distribution can be approximated by a Gaussian function, which is well localized around its average value. In this case the majority of rod responses will be "concentrated" around the saturation amplitude A_s , see Fig. 6.7a (black bars). For the thermal source the photon probability distribution is a decaying exponential function, and even if the average number of photons is large, there is always a non-vanishing probability for emission of pulses with just a few photons. Hence some responses will have amplitudes well below A_s , see Fig. 6.7b (black bars). Because of contribution of such events, the average amplitude becomes smaller and as a result, the cell response is less steep. Remarkably, obtained results are consistent with earlier findings of frequency of seeing experiments by Teich and colleagues, suggesting that responses of isolated rod cells are at the basis of those effects [85–87].

Saturation also reveals itself in dependencies of the signal-to-noise ratio (SNR), see Fig. 6.8. For stimulation with weak pulses, the cell response is linear and it reflects the statistics of the light source: $\text{Var} A = \langle A \rangle \rightarrow \text{SNR} = \sqrt{\langle A \rangle}$ for the coherent source, and $\text{Var} A = \langle A \rangle^2 \rightarrow \text{SNR} = 1$ for the thermal source. In the saturation regime the cell most likely to return responses with amplitudes equal to A_s , which leads to reduction of the variance and growth of the SNR for both sources. This is confirmed by numerical simulations, see lines in Fig. 6.8.

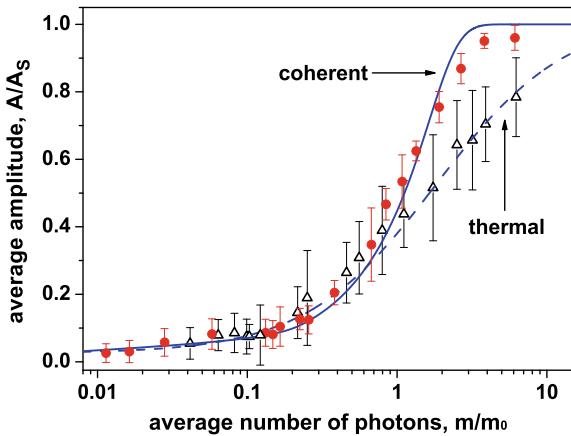


Fig. 6.6 Dependence of the average amplitude of the rod response on the average number of photons for coherent (red circles, solid line) and thermal (black triangles, dashed line) sources. Amplitudes are normalized for a saturation amplitude A_s , and the number of photons is normalized to the number of photons in the coherent pulse m_0 which initiated a response of a half saturation amplitude ($m_0 \approx 550\text{--}2500$ photon per pulse (number of photons is not adjusted for the efficiency of photon delivery to the cell)). The curves are averaged over 6 cells from 6 different animals and error bars show \pm s.d. Lines show numerical calculations according to the photon counting model. The dependencies show different trends due to differences in photon statistics of the light sources. The difference becomes especially noticeable when the response is close to the saturation. From [19]

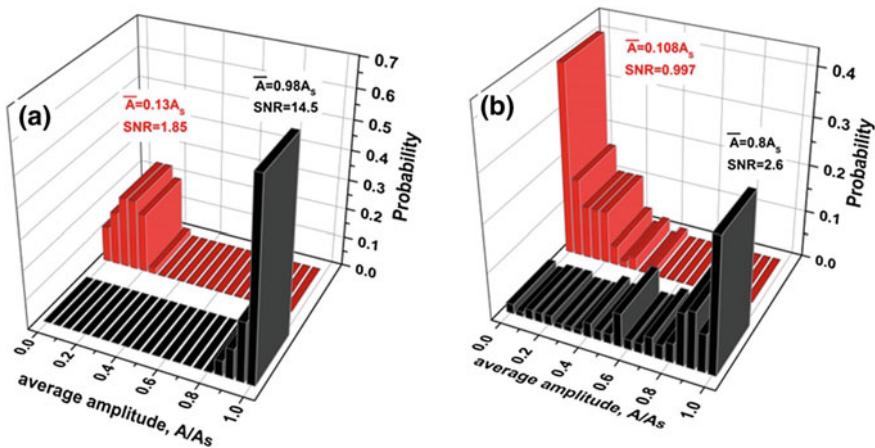


Fig. 6.7 Probability histograms of normalized amplitudes measured for coherent (a) and pseudothermal (b) sources at different values of the average amplitude of cell response (normalized to the saturation amplitude A_s), shown with the corresponding SNR values. Red and black bars correspond to linear and close-to-saturation regime of cell response, respectively. Adapted from [19]

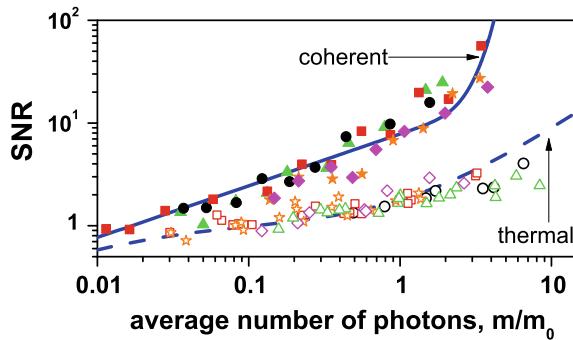


Fig. 6.8 Dependence of the signal-to-noise ratio (SNR) on the average number of photons normalised to the number of photons in the coherent pulse m_0 which initiated a response of a half saturation amplitude. Closed symbols, solid line show the results for the coherent source. Open symbols, dashed line show results for the thermal source. Different symbol shapes correspond to data obtained from different cells. Lines show numerical calculations according to the photon counting model. Adapted from [19]

It is also interesting to analyse the trend of the second order correlation function $g^{(2)}$, see Fig. 6.9. When the rod cell is stimulated by a coherent source $g^{(2)}$ is equal to 1 and it does not depend on the number of photons. This result is expected because after the beam-splitter photons impinging on the cell and on the APD are not correlated [96]. For the thermal source $g^{(2)}$ decreases with the increase of the number of photons from theoretical value of 2 to 1. In the saturation regime the cell consistently produces responses with amplitudes equal to A_s . The response amplitude can be traced out of the averaging $\langle AK \rangle \approx A_s \langle K \rangle$ in the nominator of $g^{(2)}$, which results in the asymptotic value of $g^{(2)} = 1$.

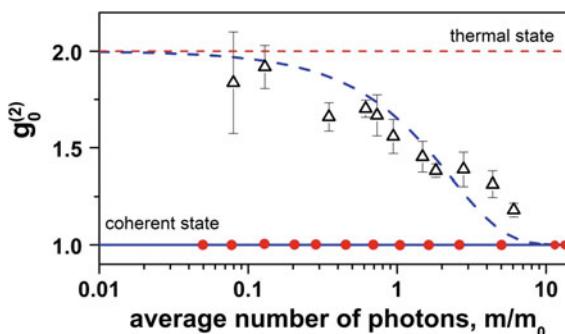


Fig. 6.9 Dependence of the second order correlation function $g_0^{(2)}$ on the average number of photons for coherent (red circles, solid line) and thermal (black triangles, dashed line) sources. Lines show numerical calculations according to the photon counting model. Saturation of the cell response leads to the decrease of $g_0^{(2)}$ for the thermal source, while $g_0^{(2)}$ stays constant for the coherent source. Error bars show \pm s.d. Adapted from [19]

6.3.3 Interfacing Rod Cells with a True Single Photon Source

Behavioural and electrophysiological experiments evidenced that rod cells are capable to respond to light at the level of single photons [25, 29]. In those studies conventional light sources such as lamps, lasers, and LEDs were used. As discussed above, such sources exhibit fundamental fluctuations in the number of emitted photons. These fluctuations become crucial at ultra-low intensities of light pulses. In this case, variance of the number of emitted photons becomes comparable or even larger than the mean value. Hence, the single photon sensitivity of rod cells can only be inferred from the statistical analysis. Moreover, the uncertainty in the number of photons, impinging on the cell, hinders accurate characterization of underlying bio-chemical mechanisms of the photo-transduction. The intrinsic physiological noise of the cell cannot be distinguished from the noise due to fluctuations of light stimuli. A feasible way to address these problems is to use a specially engineered light source, which produces pulses with given number of photons down to single photon level. This experiment was recently realized in the work of Phan and colleagues [95].

Considerable interest is focused on the development of single photon sources for the purposes of secure communication and quantum computing. Phan and colleagues, engineered a single photon source which allows studying interaction of single photons with a biological object. The single photon source is based on a process of spontaneous parametric down conversion (SPDC) [98]. In the SPDC a fraction of a laser pulse (pump), propagating in a non-linear optical crystal, is converted into a pair of photons (signal and idler), which obey conservation of energy and momentum:

$$\omega_p = \omega_s + \omega_i \quad k_p = k_s + k_i,$$

where $\omega_{p,s,i}$ and $k_{p,s,i}$ are frequencies and wave vectors of pump, signal, and idler photons, respectively. SPDC is a probabilistic process, but conservation laws guarantee that the signal and idler photons are emitted simultaneously and in strictly defined directions. In the experiment, signal and idler photons were created from a pulsed UV-laser with $\lambda_p = 266$ nm. The photons were of the same wavelengths of $\lambda_s = \lambda_i = 532$ nm, which were chosen to maximize the absorption by the cell. The signal and idler photons travelled in two different directions, which form an angle of $\pm 3^\circ$ to the direction of the pump beam, see Fig. 6.10. The pump power was adjusted to minimise the chance of creation of more than one pair from a single pump pulse.

Correlation between signal and idler photons can be readily exploited in generation of single photon pulses. A single-photon avalanche photodiode (APD) was put in the signal beam, and a fast optical shutter (acousto-optical modulator) in the idler beam, see Fig. 6.10. The APD output was used to trigger opening of the shutter in the idler beam. Once the signal photon was detected by the APD,

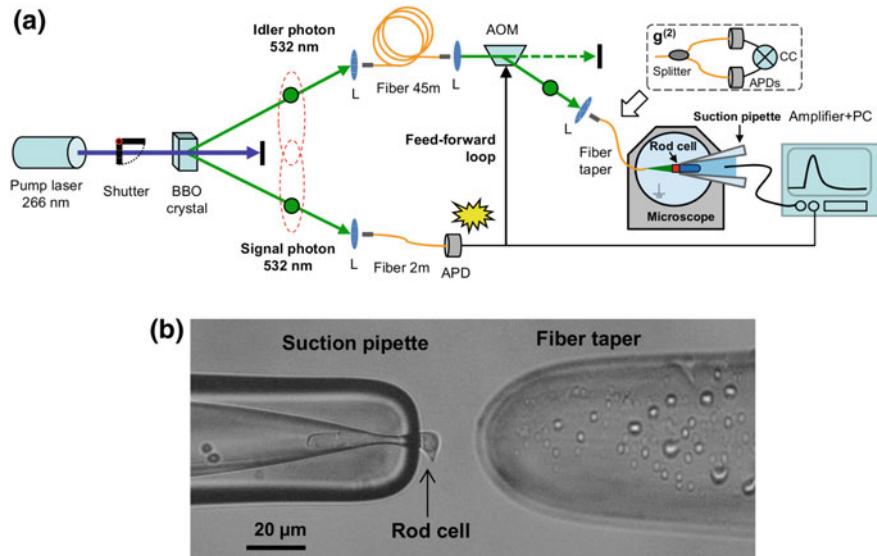


Fig. 6.10 **a** Experimental setup for single photon stimulation of retinal rod cells. A pulsed laser at 266 nm pumps a Beta barium borate (BBO) crystal, where spontaneous parametric down conversion (SPDC) occurs. Signal and idler photons are emitted at 532 nm. Signal photon is detected by the avalanche photodiode (APD) and it triggers an acousto-optical modulator (AOM) in the idler arm. The idler photon travels in a 45 m long fiber and gets deflected by the AOM. It is then addressed to a tapered fiber, pointing at the rod cell. Excitation of the rod cell is recorded with the suction electrode technique. Inset shows a setup used for the measurement of the correlation function of the single photon consisting of a 50/50 beam splitter and two APDs connected to a coincidence circuit. **b** Microscope image shows position of the rod cell in the recording pipette interfaced with a taper of the optical fiber. Adapted from [95]

the shutter opened for a brief period of time, during which the idler photon passed through it. The idler photon was delayed by an optical fiber, to compensate for electronic delays, and a time jitter of the shutter. In case if the APD did not detect a signal photon the shutter remained closed and no light pulse was sent to the cell. Thus, observation of a photocount of the APD in the signal beam heralded presence of a single photon in the idler beam. The heralded single photon was then directed to the rod cell in the recording pipette via an optical fiber [99].

Single photon sources are conventionally assessed with the second order intensity correlation function $g^{(2)}$. In the experiment $g^{(2)}$ of heralded photons was measured using a 50/50 beamsplitter and two APDs with their outputs connected to a coincidence circuit, see inset Fig. 6.10. The source yielded $g^{(2)} = 0.08 \pm 0.06$, which means that the probability of obtaining a multiphoton event is 12.5 times less, compared to the coherent source with $g^{(2)} = 1$.

Rod cells were isolated from dark adapted retinas of frogs *Xenopus laevis*. Real-time monitoring of the membrane current was realized using the suction electrode technique [73, 74]. Waveforms, accompanied by the APD photocounts,

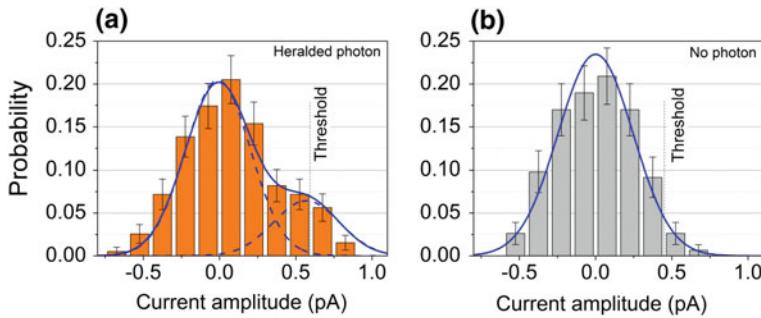


Fig. 6.11 **a** Probability distribution of cell electric response amplitudes when the APD heralds a presence of a single photon and **b** when there is no heralding signal (dark noise). Solid lines are Gaussian fits. The vertical dashed lines indicate the threshold level ($A \geq 0.45$ pA) for calculation of single-photon responses. Error bars show \pm s.d. Adapted from [95]

were used to analyse single photon responses. Waveforms, acquired in the absence of the APD photocounts, are used to analyse the dark noise.

Probability histogram of waveform amplitudes for the case, when the APD heralds the presence of a single photon, is shown Fig. 6.11a (orange bars). It shows two distinctive peaks: a non-response peak (centered about 0 pA) and a single photon response peak (centered about 0.6 pA). Histograms of the dark noise for the same cell is plotted in Fig. 6.11b (grey bars). It reveals few *single photon-like* dark noise responses, which are less frequent than in case of single photon stimulation.

From the amplitude histogram, a criterion-based method is established to identify single-photon responses. Waveforms with amplitudes higher than the criterion level are categorized as “*single photon responses*”, and lower than the criterion level as “*no responses*”. Based on discretisation of the histogram in Fig. 6.11a, the criterion level is selected at 0.45 pA.

Probability of occurrence of single photon responses, which satisfy the above mentioned criteria, is consistently higher for single photon stimulation, than for the dark noise, see Fig. 6.12. Thus responsiveness of the cell to stimuli, produced by the heralded single photon source is clearly justified.

Averaged waveform of single photon responses is shown in Fig. 6.13a. It shows consistent amplitude, time-to-peak, and duration. Averaged zero-photon response does not show any distinctive shape and appear almost flat, see Fig. 6.13b. The parameters of the waveforms for ten different cells from ten different animals, are summarized in Table 2. Cell-to-cell variations could be attributed to differences in intrinsic capabilities of single photon detection, different resistances of the suction electrodes, and different efficiencies of light coupling. Nevertheless, the results allow clear identification of amplitudes of single photon responses for the majority of studied cells within the range 0.5–0.7 pA.

Quantum efficiency (QE) of the rod cell is an important parameter, which characterizes its ability to respond to light. QE can be defined similar to man-made single photon detectors as a coefficient between the number of responses and the

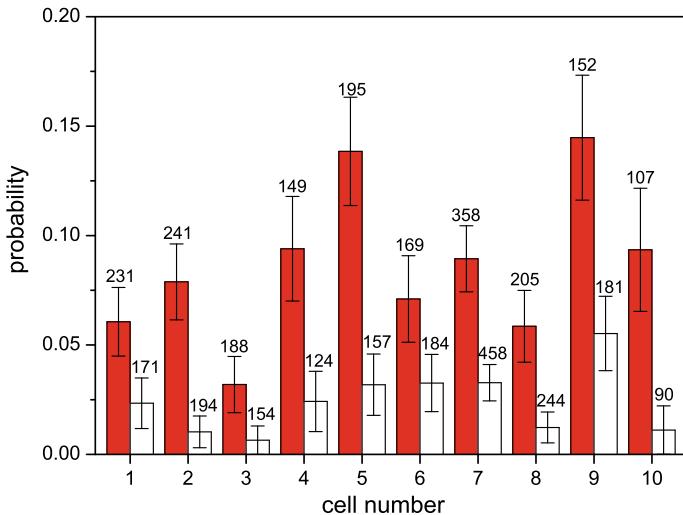


Fig. 6.12 Probability of the occurrence of single-photon responses, satisfying the criterion ($A \geq 0.45$ pA), when the APD heralds a single photon (red bars) and for the dark noise (white bars). Labels correspond to the number of experimental trials. Error bars show \pm s.d. Adapted from [95]

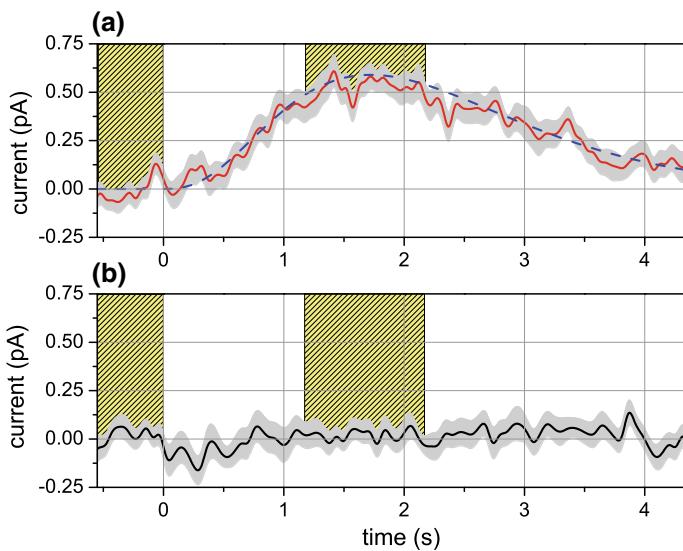


Fig. 6.13 **a** Representative average waveform of the cell responses (cell # 5 in Fig. 6.12) to single photons (red solid line), and **b** of non-responses (bandwidth 20 Hz, 27 traces). Blue dashed line in (a) is a theoretical curve based on the impulse response of the Poisson filter. Shutter in the pump beam is opened for a period of 100 ms at $t = 0$. Yellow shaded regions indicate widths of time windows for calculation of waveform amplitudes. Grey shaded regions in (a, b) show \pm s.e.m. Adapted from [95]

Table 2 Parameters of single photon responses for rod cells from dark adapted retinas of frogs *Xenopus laevis** [95]

Cell #	Amplitude, pA	Time to peak, s	Full width at the level of half amplitude, s	Number of traces
1	0.54 ± 0.02	2.7 ± 0.2	2.8 ± 0.1	14
2	0.58 ± 0.02	1.8 ± 0.1	2.2 ± 0.1	19
3	0.59 ± 0.03	2.9 ± 0.2	3.6 ± 0.2	6
4	0.57 ± 0.02	0.7 ± 0.1	1.0 ± 0.1	14
5	0.65 ± 0.03	1.7 ± 0.1	2.5 ± 0.1	27
6	0.57 ± 0.03	1.9 ± 0.1	2.3 ± 0.2	12
7	0.59 ± 0.02	2.8 ± 0.1	2.4 ± 0.15	32
8	0.60 ± 0.05	1.8 ± 0.1	2.0 ± 0.1	12
9	0.65 ± 0.03	0.7 ± 0.1	1.6 ± 0.1	22
10	0.60 ± 0.04	0.95 ± 0.15	1.3 ± 0.1	10

*Values are the mean \pm s.e.m.

number of impinging photons. Let us consider that N photon pairs were emitted from the SPDC crystal. The number of photons, detected by the APD in the signal beam, is $n_{\text{APD}} = \eta_{\text{APD}}N$, where η_{APD} is the APD quantum efficiency. After that, n_{APD} trigger pulses will be sent to the optical shutter in the idler arm. The corresponding number of photons detected by the cell is $n_{\text{Cell}} = \eta_{\text{Cell}}\eta_{\text{APD}}$, where η_{Cell} is the unknown quantum efficiency of the cell (here, for the moment, we neglect optical losses in the idler beam). From the above η_{Cell} can be found as a ratio of the two directly measured experimental values, $\eta_{\text{Cell}} = n_{\text{Cell}}/n_{\text{APD}}$. In contrast to conventional techniques of QE calibration, this method does not require the use of any pre-calibrated devices [100]. From data in Fig. 6.12, and taking into account optical losses in the idler channel, and the dark noise, the average value of the QE for ten studied cells is $\eta_{\text{Cell}} = 29 \pm 4.7\%$. This value is in good agreement with earlier physiological studies for *Xenopus*.

The experiment with a true single photon source provided a conclusive proof of single photon sensitivity of rod cells. It also demonstrated application of the reference-free calibration technique to measurement of the cell's QE.

6.4 Conclusions and Future Perspectives

Fast progress in molecular biology and protein chemistry offers new ways of measuring the voltage output from rod cells and novel means to manipulate the phototransduction mechanisms. Recent development of voltage-sensitive fluorescent proteins and dyes (reviewed in: [101]) could be useful for measurements of membrane potential in rod cells under illumination. Potentially the red or near infra-red fluorescent proteins [102], which do not overlap in their excitation and

emission properties with spectral sensitivity of rod cells and fused with voltage-sensitive sensor might be helpful in elucidating the distribution of membrane potential in rod cells under dark conditions and after pulse of photons. The other already used means is to modify or correct elements of phototransduction cascades. Recently adeno-associated viral delivery of genes proved to be effective for functional expression of proteins in human and mice retinal photoreceptors [103].

From the optical perspective, an interesting development could be studies of interaction of rod cells with pulses carrying precise number of photons (more than one). Following the proposal by Teich and colleagues, such an experiment would allow characterizing statistical properties of distinctive steps in the phototransduction in much more precise way, as compared to Poissonian light sources [85–87]. It may also lead to direct observation of quantum properties, such as entanglement and squeezing, possibly with naked eyes [104].

References

1. J.E. Dowling, *The Retina: An Approachable Part of the Brain* (Belknap Press of Harvard University Press, Cambridge, MA, 2012). Revised Edition
2. C.R. Braekevelt, S.A. Smith, B.J. Smith, Fine structure of the retinal photoreceptors of the barred owl (*Strix varia*). *Histol. Histopathol.* **11**(1), 79–88 (1996)
3. M. Joseph, A. Corless, Minimum diameter limit for retinal rod outer segment disks. *Development of Order in the Visual System*, ed. by S.R. Hilfer et al. (Springer, New York Inc., 1986), pp. 127–142
4. K. Palczewski, G protein-coupled receptor rhodopsin. *Annu. Rev. Biochem.* **75**, 743–767 (2006)
5. M.L. Woodruff, M.D. Bownds, Amplitude, kinetics, and reversibility of a light-induced decrease in guanosine 3',5'-cyclic monophosphate in frog photoreceptor membranes. *J. Gen. Physiol.* **73**(5), 629–653 (1979)
6. H.W. Choe, Y.J. Kim, J.H. Park, T. Morizumi, E.F. Pai, N. Krauss, K.P. Hofmann, P. Scheerer, O.P. Ernst, Crystal structure of metarhodopsin II. *Nature* **471**, 651–655 (2011)
7. K. Palczewski, T. Kumazaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, I. Le Trong, D.C. Teller, T. Okada, R.E. Stenkamp et al., Crystal structure of rhodopsin: a G protein-coupled receptor. *Science* **289**, 739–745 (2000)
8. L. Stryer, Exploring light and life. *J. Biol. Chem.* **287**, 15164–15173 (2012)
9. T.D. Lamb, Gain and kinetics of activation in the G-protein cascade of phototransduction. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 566–570 (1996)
10. T.D. Lamb, E.N. Pugh Jr., Phototransduction, dark adaptation, and rhodopsin regeneration the proctor lecture. *Invest. Ophthalmol. Vis. Sci.* **47**, 5138–5152 (2006)
11. E.N. Pugh Jr., T.D. Lamb, Cyclic GMP and calcium: the internal messengers of excitation and adaptation in vertebrate photoreceptors. *Vis. Res.* **30**, 1923–1948 (1990)
12. R.H. Cote, M.A. Brunnock, Intracellular cGMP concentration in rod photoreceptors is regulated by binding to high and moderate affinity cGMP binding sites. *Biol. Chem.* **268** (23), 17190–17198 (1993)
13. X. Zhang, R.H. Cote, cGMP signaling in vertebrate retinal photoreceptor cells. *Front Biosci.* **10**, 1191–1204 (2005)

14. E.E. Fesenko, S.S. Kolesnikov, A.L. Lyubarsky, Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. *Nature* **313**, 310–313 (1985)
15. W.H. Cobbs, E.N. Pugh Jr., Cyclic GMP can increase rod outer-segment light-sensitive current 10-fold without delay of excitation. *Nature* **313**, 585–587 (1985)
16. K. Matulef, W.N. Zagotta, Cyclic nucleotide-gated ion channels. *Annu. Rev. Cell Dev. Biol.* **19**, 23–44 (2003)
17. R.R. Birge, Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin. *Biochim. Biophys. Acta* **1016**, 293–327 (1990)
18. B.K.-K. Fung, J.B. Hurley, L. Stryer, Flow of information in the light-triggered cyclic nucleotide cascade of vision. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 152–156 (1981)
19. N. Sim, M.F. Cheng, D. Bessarab, C.M. Jones, L.A. Krivitsky, Measurement of photon statistics with live photoreceptor cells. *Phys. Rev. Lett.* **109**, 113601 (2012)
20. M.E. Burns, E.N. Pugh Jr., Lessons from photoreceptors: turning off G-protein signaling in living cells. *Physiology (Bethesda)* **25**, 72–84 (2010)
21. C.M. Krispel, D. Chen, N. Melling, Y.J. Chen, K.A. Martemyanov, N. Quillinan, V.Y. Arshavsky, T.G. Wensel, C.K. Chen, M.E. Burns, RGS expression rate-limits recovery of rod photoresponses. *Neuron* **51**, 409–416 (2006)
22. C.K. Chen, M.L. Woodruff, F.S. Chen, D. Chen, G.L. Fain, Background light produces a recoverin-dependent modulation of activated-rhodopsin lifetime in mouse rods. *J. Neurosci.* **30**, 1213–1220 (2010)
23. W.H. Cobbs, E.N. Pugh Jr., Kinetics and components of the flash photocurrent of isolated retinal rods of the larval salamander, *Ambystoma tigrinum*. *J. Physiol.* **394**, 529–572 (1987)
24. P. Bisegna, G. Caruso, D. Andreucci, L. Shen, V.V. Gurevich, H.E. Hamm, E. DiBenedetto, Diffusion of the second messengers in the cytoplasm acts as a variability suppressor of the single photon response in vertebrate phototransduction. *Biophys. J.* **94**, 3363–3383 (2008)
25. F. Rieke, D.A. Baylor, Origin of reproducibility in the responses of retinal rods to single photons. *Biophys. J.* **75**, 1836–1857 (1998)
26. U.B. Kaupp, R. Seifert, Cyclic nucleotide-gated ion channels. *Physiol. Rev.* **82**(3), 769–824 (2002)
27. E. Eismann, F. Müller, S.H. Heinemann, U.B. Kaupp, A single negative charge within the pore region of a cGMP-gated channel controls rectification, Ca²⁺ blockage, and ionic selectivity. *Proc. Natl. Acad. Sci. U.S.A.* **91**(3), 1109–1113 (1994)
28. K.W. Yau, D.A. Baylor, Cyclic GMP activated conductance of retinal photoreceptor cells. *Annu. Rev. Neurosci.* **12**, 289–327 (1989)
29. F. Rieke, D.A. Baylor, Single photon detection by rod cells of the retina. *Rev. Mod. Phys.* **70**, 1027–1036 (1998)
30. T. Doan, A. Mendez, P.B. Detwiler, J. Chen, F. Rieke, Multiple phosphorylation sites confer reproducibility of the rod's single-photon responses. *Science* **313**, 530–533 (2006). PMID: 16873665, <http://dx.doi.org/10.1126/science.1126612>
31. A.W. Azevedo, T. Doan, H. Moaven, I. Sokal, F. Baameur, S.A. Vishnivetskiy, K.T. Homan, J.J. Tesmer, V.V. Gurevich, J. Chen, F. Rieke, C-terminal threonines and serines play distinct roles in the desensitization of rhodopsin, a G protein-coupled receptor. *Elife* **4** (2015). <https://doi.org/10.7554/elife.05981>
32. V. Torre, H.R. Matthews, T.D. Lamb, Role of calcium in regulating the cyclic GMP cascade of phototransduction in retinal rods. *Proc. Natl. Acad. Sci. U.S.A.* **83**(18), 7109–7113 (1986)
33. M. Capovilla, L. Cervetto, V. Torre, The effect of phosphodiesterase inhibitors on the electrical activity of toad rods. *J. Physiol.* **343**, 277–294 (1983)
34. M. Capovilla, L. Cervetto, V. Torre, Effects of changing external potassium and chloride concentrations on the photoresponses of *Bufo bufo* rods. *J. Physiol.* **307**, 529–551 (1980)
35. E.N. Pugh Jr., T.D. Lamb, Amplification and kinetics of the activation steps in phototransduction. *Biochim. Biophys. Acta* **1141**(2–3), 111–149 (1993)
36. D.A. Baylor, B.J. Nunn, J.L. Schnapf, The photocurrent, noise and spectral sensitivity of rods of the monkey *Macaca fascicularis*. *J. Physiol.* **357**, 575–607 (1984)

37. S. Asteriti, S. Grillner, L. Cangiano, A Cambrian origin for vertebrate rods. *eLife* **4**, e07166 (2015). <https://doi.org/10.7554/elife.07166>
38. J.R. Sanes, S.L. Zipursky, Design principles of insect and vertebrate visual systems. *Neuron* **66**(1), 15–36 (2010). <https://doi.org/10.1016/j.neuron.2010.01.018>
39. C. Montell, Visual transduction in *Drosophila*. *Annu. Rev. Cell Dev. Biol.* **15**, 231–268 (1999)
40. C. Montell, *Drosophila* visual transduction. *Trends Neurosci.* **35**, 356–363 (2012)
41. R.C. Hardie, M. Juusola, Phototransduction in *Drosophila*. *Curr. Opin. Neurobiol.* **34C**, 37–45 (2015)
42. R.C. Hardie, Phototransduction in *Drosophila melanogaster*. *J. Exp. Biol.* **204**(Pt 20), 3403–3409 (2001)
43. A. Auerbach, F. Sachs, Flickering of a nicotinic ion channel to a subconductance state. *Biophys. J.* **42**(1), 1–10 (1983)
44. O. Alvarez, C. Gonzalez, R. Latorre, Counting channels: a tutorial guide on ion channel fluctuation analysis. *Adv. Physiol. Educ.* **26**(1–4), 327–341 (2002)
45. I. Lestas, G. Vinnicombe, J. Paulsson, Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–178 (2010)
46. D.G. Spiller, C.D. Wood, D.A. Rand, M.R. White, Measurement of single-cell dynamics. *Nature* **465**, 736–745 (2010)
47. P.N. Steinmetz, R.L. Winslow, Optimal detection of flash intensity differences using rod photocurrent observations. *Neural Comput.* **11**(5), 1097–1111 (1999)
48. R.A. Yotter, D.M. Wilson, A review of photodetectors for sensing light-emitting reporters in biological systems. *IEEE Sens. J.* **3**, 288–303 (2003)
49. G.N. Gol'tsman, O. Okuney, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, Picosecond superconducting single-photon optical detector. *Appl. Phys. Lett.* **79**(6), 705–707 (2001)
50. M. Dandin, P. Abshire, High signal-to-noise ratio avalanche photodiodes with perimeter field gate and active readout. *IEEE Electron Device Lett.* **33**(4), 570–572 (2012)
51. K. Kolb, Signal-to-noise ratio of Geiger-mode avalanche photodiode single-photon counting detectors. *Opt. Eng.* **53**(8), 081904 (2014)
52. J.E. Sulston, E. Schierenberg, J.G. White, J.N. Thomson, The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**(1), 64–119 (1983)
53. A.S. Chiang, C.Y. Lin, C.C. Chuang, H.M. Chang, C.H. Hsieh, C.W. Yeh, C.T. Shih, J. J. Wu, G.T. Wang, Y.C. Chen, C.C. Wu, G.Y. Chen, Y.T. Ching, P.C. Lee, C.Y. Lin, H.H. Lin, C.C. Wu, H.W. Hsu, Y.A. Huang, J.Y. Chen, H.J. Chiang, C.F. Lu, R.F. Ni, C.Y. Yeh, J.K. Hwang, Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol.* **21**(1), 1–11 (2011)
54. L. Chittka, J. Niven, Are bigger brains better? *Curr. Biol.* **19**, R995–R1008 (2009)
55. S. Herculano-Houzel, The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc. Natl. Acad. Sci. U.S.A.* **109**(Suppl 1), 10661–10668 (2012)
56. S.M. Wu, Synaptic transmission in the outer retina. *Annu. Rev. Physiol.* **56**, 141–168 (1994)
57. W. Bialek, W.G. Owen, Temporal filtering in retinal bipolar cells. Elements of an optimal computation? *Biophys. J.* **58**(5), 1227–1233 (1990)
58. H. Markram, E. Müller, S. Ramaswamy, M.W. Reimann, M. Abdellah, C.A. Sanchez, A. Ailamaki, L. Alonso-Nanclares, N. Antille, S. Arsever, G.A. Kahou, T.K. Berger, A. Bilgili, N. Buncic, A. Chalimourda, G. Chindemi, J.D. Courcol, F. Delalondre, V. Delattre, S. Druckmann, R. Dumusc, J. Dynes, S. Eilemann, E. Gal, M.E. Gevaert, J.P. Ghobril, A. Gidon, J.W. Graham, A. Gupta, V. Haenel, E. Hay, T. Heinis, J.B. Hernando, M. Hines, L. Kanari, D. Keller, J. Kenyon, G. Khazen, Y. Kim, J.G. King, Z. Kisvarday, P. Kumbhar, S. Lasserre, J.V. Le Bé, B.R. Magalhães, A. Merchán-Pérez, J. Meystre, B.R. Morrice, J. Müller, A. Muñoz-Céspedes, S. Muralidhar, K. Muthurasa, D. Nachbaur, T.H. Newton, M. Nolte, A. Ovcharenko, J. Palacios, L. Pastor, R. Perin, R. Ranjan, I. Riachi, J.R. Rodríguez, J.L. Riquelme, C. Rössert, K. Sfyarakis, Y. Shi, J.C. Shillecock, G. Silberberg, R.

- Silva, F. Tauheed, M. Telefont, M. Toledo-Rodriguez, T. Tränkler, W. Van Geit, J.V. Díaz, R. Walker, Y. Wang, S.M. Zaninetta, J. DeFelipe, S.L. Hill, I. Segev, F. Schürmann, Reconstruction and simulation of neocortical microcircuitry. *Cell* **163**(2), 456–492 (2015)
59. R.G. Boothe, *Perception of the Visual Environment*. Psychology (Springer Science & Business Media, 2001), 408 pages. ISBN: 978-0-387-98790-3 (Print) 978-0-387-21650-8 (Online)
60. A.P. Sampath, F. Rieke, Selective transmission of single photon responses by saturation at the rod-to-rod bipolar synapse. *Neuron* **41**(3), 431–443 (2004)
61. D. Attwell, S. Borges, S.M. Wu, M. Wilson, Signal clipping by the rod output synapse. *Nature* **328**(6130), 522–524 (1987)
62. S. Barnes, V. Merchant, F. Mahmud, Modulation of transmission gain by protons at the photoreceptor output synapse. *Proc. Natl. Acad. Sci. U.S.A.* **90**(21), 10081–10085 (1993)
63. A.J. Mercer, W.B. Thoreson, The dynamic architecture of photoreceptor ribbon synapses: cytoskeletal, extracellular matrix, and intramembrane proteins. *Vis. Neurosci.* **28**(6), 453–471 (2011)
64. A. Bharioke, D.B. Chklovskii, Automatic adaptation to fast input changes in a time-invariant neural circuit. *PLoS Comput. Biol.* **11**(8), e1004315 (2015). <https://doi.org/10.1371/journal.pcbi.1004315>
65. S.P. Langley, The bolometer and radiant energy, in *Proceedings of the American Academy of Arts and Science*, vol. 16 (American Academy of Arts & Sciences, May 1880–Jun 1881), pp. 342–358. <https://doi.org/10.2307/25138616>, <http://www.jstor.org/stable/25138616>
66. J. von Kries, J.A.E. Eyster, Über die zur Erregung des Sehorgans efforderlichen Energiemengen. *Z. Sinnesphysiol.* **41**, 373–394 (1907)
67. A. Verkhratsky, O.A. Krishtal, O.H. Petersen, From Galvani to patch clamp: the development of electrophysiology. *Pflugers Arch.* **453**(3), 233–247 (2006)
68. T. Tomita, A. Funaishi, Studies on intraretinal action potential with low-resistance microelectrode. *J. Neurophysiol.* **15**(1), 75–84 (1952)
69. G.S. Brindley, Responses to illumination recorded by microelectrodes from the frog's retina. *J. Physiol.* **134**(2), 360–384 (1956)
70. A.L. Byzov, Functional properties of different cells in the retina of cold-blooded vertebrates. *Cold Spring Harb. Symp. Quant. Biol.* **30**, 547–558 (1965)
71. S.R. Grabowski, L.H. Pinto, W.L. Pak, Adaptation in retinal rods of axolotl: intracellular recordings. *Science* **176**(4040), 1240–1243 (1972)
72. R.D. Penn, W.A. Hagins, Signal transmission along retinal rods and the origin of the electroretinographic a-wave. *Nature* **223**(5202), 201–204 (1969)
73. D.A. Baylor, T.D. Lamb, K.W. Yau, The membrane current of single rod outer segments. *J. Physiol.* **288**, 589–611 (1979)
74. D.A. Baylor, T.D. Lamb, K.W. Yau, Responses of retinal rods to single photons. *J. Physiol.* **288**, 613–634 (1979)
75. R.B. Barnes, M. Czerny, Läßt sich ein Schroteffekt der Photonen mit dem Auge beobachten? *Zeitschrift für Physik* **79**(7), 436–449 (1932)
76. S. Hecht, S. Shlaer, M.H. Pirenne, Energy, quanta, and vision. *J. Gen. Physiol.* **25**(6), 819–840 (1942)
77. E. Brumberg, S. Vavilov, Visuelle Messungen der statistischen Photonenschwankungen. *Bull. Acad. Sci. U.R.S.S.* **7**, 919–941 (1933)
78. E.M. Brumberg, S.I. Vavilov, Z.M. Sverdlov, Visual measurements of quantum fluctuations. I. The threshold of vision as compared with the results of fluctuation measurements. *J. Phys.* **7**(1), 1–8 (1943)
79. S.I. Vavilov, T.V. Timofeeva, Visual measurements of quantum fluctuations. II. Fluctuations when the eye is light-adapted. *J. Phys.* **7**(1), 9–11 (1943)
80. S.I. Vavilov, T.V. Timofeeva, Visual measurements of quantum fluctuations. III. The dependence of the visual fluctuations on the wave-length. *J. Phys.* **7**(1), 12–17 (1943)
81. S.I. Vavilov, *The Microstructure of Light* (Academy of Sciences, Moscow, 1950), p. 198. (in Russian)

82. S. Hecht, S. Shlaer, M.H. Pirenne, Energy at the threshold of vision. *Science* **93**(2425), 585–587 (1941)
83. R. Gunter, The absolute threshold for vision in the cat. *J. Physiol.* **114**(1–2), 8–15 (1951)
84. S. Hecht, M.H. Pirenne, The sensibility of the nocturnal long-eared owl in the spectrum. *J. Gen. Physiol.* **23**(6), 709–717 (1940)
85. M.C. Teich, P.R. Prucnal, G. Vannucci, M.E. Breton, W.J. McGill, Multiplication noise in the human visual system at threshold: 1. Quantum fluctuations and minimum detectable energy. *J. Opt. Soc. Am.* **72**, 419–431 (1982)
86. P.R. Prucnal, M.C. Teich, Multiplication noise in the human visual system at threshold: 2. Probit estimation of parameters. *Biol. Cybern.* **43**, 87–96 (1982)
87. M.C. Teich, P.R. Prucnal, G. Vannucci, M.E. Breton, W.J. McGill, Multiplication noise in the human visual system at threshold: 3. The role of non-poisson quantum fluctuations. *Biol. Cybern.* **44**, 157–165 (1982)
88. K.W. Yau, T.D. Lamb, D.A. Baylor, Light-induced fluctuations in membrane current of single toad rod outer segments. *Nature* **269**(5623), 78–80 (1977)
89. P.B. Detwiler, J.D. Conner, R.D. Bodoia, Gigaseal patch clamp recordings from outer segments of intact retinal rods. *Nature* **300**(5887), 59–61 (1982)
90. R.D. Bodoia, P.B. Detwiler, Patch-clamp recordings of the light-sensitive dark noise in retinal rods from the lizard and frog. *J. Physiol.* **367**, 183–216 (1985)
91. J. Toyoda, H. Hashimoto, H. Anno, T. Tomita, The rod response in the frog and studies by intracellular recording. *Vis. Res.* **10**(11), 1093–1100 (1970)
92. T. Tomita, Electrical activity of vertebrate photoreceptors. *Q. Rev. Biophys.* **3**(2), 179–222 (1970)
93. J.E. Brown, L.H. Pinto, Ionic mechanism for the photoreceptor potential of the retina of *Bufo marinus*. *J. Physiol.* **236**(3), 575–591 (1974)
94. R.R. Birge, R.B. Barlow, On the molecular origins of thermal noise in vertebrate and invertebrate photoreceptors. *Biophys. Chem.* **55**, 115–126 (1995)
95. N.M. Phan, M.F. Cheng, D.A. Bessarab, L.A. Krivitsky, Interaction of fixed number of photons with retinal rod cells. *Phys. Rev. Lett.* **112**, 213601 (2014)
96. L. Mandel, E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England, 1995)
97. F.T. Arecchi, Measurement of the statistical distribution of Gaussian and laser sources. *Phys. Rev. Lett.* **15**, 912 (1965)
98. D.N. Klyshko, *Physical Foundations of Quantum Electronics* (World Scientific, Singapore, 2011)
99. N. Sim, D. Bessarab, C.M. Jones, L. Krivitsky, Method of targeted delivery of laser beam to isolated retinal rods by fiber optics. *Biomed. Opt. Express* **2**, 2926–2933 (2011)
100. A.A. Malygin, A.N. Penin, A.V. Sergienko, Absolute calibration of the sensitivity of photodetectors using a biphotonic field. *Sov. Phys. JETP Lett.* **33**, 477–481 (1981)
101. H. Mutoh, W. Akemann, T. Knöpfel, Genetically engineered fluorescent voltage reporters. *ACS Chem. Neurosci.* **3**, 585–592 (2012)
102. K.D. Piatkevich, F.V. Subach, V.V. Verkhusha, Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem. Soc. Rev.* **42**(8), 3441–3452 (2013)
103. T. Tolmachova, O.E. Tolmachov, A.R. Barnard, S.R. de Silva, D.M. Lipinski, N.J. Walker, R.E. Maclarens, M.C. Seabra, Functional expression of Rab escort protein 1 following AAV2-mediated gene delivery in the retina of choroideremia mice and human cells ex vivo. *J. Mol. Med. (Berl.)* **91**(7), 825–837 (2013)
104. E. Pomarico, B. Sanguinetti, N. Gisin, R. Thew, H. Zbinden, G. Schreiber, A. Thomas, W. Sohler, Waveguide-based OPO source of entangled photon pairs. *New J. Phys.* **11**, 113042 (2009)
105. V. Volkov, Discovering electrophysiology in photobiology: a brief overview of several photobiological processes with an emphasis on electrophysiology. *Commun. Integr. Biol.* **7**, e28423 (2014)

Chapter 7

Controlling Light with Light via Interference on Photonic Metamaterials



Xu Fang, Kevin F. MacDonald and Nikolay I. Zheludev

Abstract It has been discovered in recent years that interactions between coherent light waves and thin-film media can, if the film is much thinner than the wavelength, lead to controllable energy exchange between incident and scattered waves and thereby to a plethora of new technological opportunities. Planar photonic metamaterials—ultrathin media with nano-engineered optical properties—can realize the full potential of this concept to change optical data processing paradigms, spectroscopy and nonlinear optics. We describe how coherent interactions in metamaterials can facilitate nonlinear light-by-light control functions with THz bandwidth at arbitrarily low intensities.

7.1 Introduction

7.1.1 Concepts and Applications of ‘Coherent Control’

Coherent control is a widely used and variously understood term, indeed coherent control concepts are found in many physical science disciplines, from chemistry to semiconductor physics, and quantum mechanics [1–3]. In photonics, broadly,

X. Fang · K. F. MacDonald (✉) · N. I. Zheludev
Optoelectronics Research Centre & Centre for Photonic Metamaterials,
University of Southampton, Highfield, Southampton SO17 1BJ, UK
e-mail: kfm@orc.soton.ac.uk

X. Fang
e-mail: x.fang@soton.ac.uk

N. I. Zheludev
e-mail: niz@orc.soton.ac.uk

N. I. Zheludev
Centre for Disruptive Photonic Technologies, School of Physical
and Mathematical Sciences & The Photonics Institute,
Nanyang Technological University, 637371 Singapore, Singapore

the term relates to the use of optical coherence to control light-matter interactions. Control schemes may invoke spatial coherence, temporal coherence, or both.

The terminology is well-understood, for example, in chemistry where it refers to the control of photo-induced chemical reactions. Before the technique was introduced, control of optically excited reactions relied predominantly on the manipulation of the frequency (wavelength) of incident light, but under these circumstances it is very difficult to isolate desired reactions from undesired reactions if their excitation energies are closely spaced. Coherent control provides for the triggering of selected reactions: typically, ultrashort laser pulses with carefully designed temporal profiles are employed to induce constructive and destructive quantum mechanical inferences for desired and undesired chemical reaction, respectively [1].

In nanophotonics, and indeed in the specific context here of ultrathin photonic metasurfaces, the same concepts of constructive and destructive interference are harnessed (albeit largely in the regime of classical electrodynamics) to manipulate local electromagnetic fields and thereby to coherently control nanoscale light-matter interactions. By controlling the spatial and/or temporal intensity and phase distribution of incident light one may, for example, dynamically control the localization of energy—the positioning of ‘hot-spots’—within a nanostructured material landscape [4–6]; and the excitation, spatial distribution and propagation of surface plasmons in nanostructured metal/dielectric environments [7–10].

7.1.2 *Coherent Light-by-Light Control in Ultrathin Media*

Here, we consider coherent control in ultrathin films and planar metasurfaces, i.e. media of deeply subwavelength thickness. We assume that they are illuminated normally by a pair of counter-propagating, collinearly polarized coherent light beams, and further that their properties are strictly linear, that they do not change the polarization state of light, and that their response to light is solely of an electric dipole nature (which is to say, as is common practice in analysing the optical response of deeply sub-wavelength structures, that we ignore magnetic and higher-order electric multipoles which respond to other field components such as magnetic field and gradient of electric field). As such, they are sensitive only to the *local* electric component of the incident electromagnetic field. So when positioned at a node of the standing wave formed by the counter-propagating incident beams (taking them to be of equal intensity), where said field is always zero, the film will not interact with the light—it will be perfectly transparent; both incident beams will propagate as if the film were not there. On the other hand, at an antinode where the local field is strong, light-matter interactions will be correspondingly strong. Indeed, as the following analysis will show, under the right circumstances the thin film may absorb ‘perfectly’ (i.e. absorb 100% of the incident light).

The optical response of the ultrathin film thus depends strongly on its position within the standing wave field, i.e. the relative phase of the two incident beams in the plane of the film. In such a system, if the position of the film is fixed, then a

change in the intensity or phase of one incident light beam will affect the other's interaction with the film, e.g. changing the level of absorption and therefore transmission/reflection. Which is to say that the coherent interaction mediated by the film provides a mechanism for controlling one beam of light with another, without recourse to any intrinsic material nonlinearity [11].

High-contrast coherent all-optical modulation, via control of resonant absorption in planar plasmonic metamaterials—nanostructured films much thinner than the wavelength of light, has been demonstrated in continuous wave and short pulse regimes down to ~6 fs (the modulation bandwidth being limited only by resonance linewidth) [11–13], and at extremely low energy levels, even down to the quantum, single-photon [14]. This coherent control paradigm can be applied to all kinds of light-matter interaction beyond absorption, as demonstrations have shown in relation to the relative efficiency of normal and anomalous refraction from phase-gradient metasurfaces [15], the manifestation of linear and circular birefringence and dichroism [16], and four-wave mixing in graphene multilayers [17]. It can also be spatially multiplexed over the plane of the film for applications such as all-optical image processing, pattern recognition and dynamic focusing [18–20]. And, moving beyond the electric dipole approximation and collinear polarization assumption introduced above, can facilitate novel modes of coherent-illumination spectroscopy, for example to selectively excite and thereby disentangle electric and magnetic resonances [21–23]. In what follows we concentrate on its application to the provision of an effective optical nonlinearity, which may be applied to a variety of all-optical signal processing functions.

7.1.3 *Effective Nonlinearity via Coherent Control*

Many of today's information and communications technologies (ICT) are existentially reliant on photonics, but for the most part these systems remain 'opaque', essentially because no optical equivalent of the transistor exists [24]. Network data is transported optically but processed electronically, with the conversions between formats limiting bandwidth while consuming energy at increasingly unsustainable rates [25]. For 90 years, since the observation of saturable absorption by Vavilov and Lewischin [26], spurred on by the invention of the laser, and driven strongly now by the performance demands of future ICT, the research community has sought the kind of fast, highly nonlinear media that would enable practically competitive all-optical switching devices, not least in terms of speed and energy-per-bit power consumption [27]. Recent works have explored gain nonlinearity in active media [28, 29], carrier-induced nonlinearity in photonic crystal cavities [30], as well as the resonant enhancement of (semiconductor, metal and low-dimensional carbon) intrinsic nonlinearities and the expression of phase-change and opto-mechanical nonlinearities in metamaterials [31–34].

Coherent control offers an entirely different approach by providing a strong effective nonlinearity, meaning simply a nonlinear relationship between the

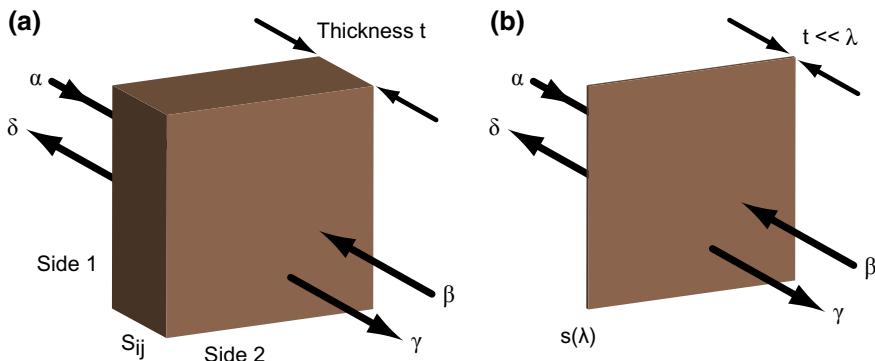


Fig. 7.1 Generic four-port coherent control device: **a** a medium, with optical properties described by a complex scattering matrix S_{ij} containing three independent parameters, is illuminated by counter-propagating coherent input light waves α and β ; waves γ and δ are the device outputs. **b** In the limit of substantially subwavelength thickness, the medium can be described by a single complex scattering coefficient $s(\lambda)$. [Reprinted with permission from ACS Photon. 4, 3000 (2017). Copyright of American Chemical Society]

intensities of signal input and output waves, based upon linear interference. The concept can be implemented at arbitrarily low intensity and does not induce harmonic distortion of the information carrier.

Consider first a device essentially as described in Sect. 7.1.2, comprising a layer of absorbing material illuminated at normal incidence by a pair of counter-propagating, collinearly polarized coherent light waves: The incident fields E_α and E_β represent input signals; the transmitted and reflected waves (fields E_γ and E_δ) propagating away from the film on either side constitute output signals (Fig. 7.1a) (Although the device has only two sides, the inputs and outputs are treated as independent channels and as such it is known as a four-port device.).

Under the above detailed assumptions of a purely linear, dipolar response with no change in the polarization state of light, the input and output fields are related by a complex scattering matrix [35] S :

$$\begin{bmatrix} E_\delta \\ E_\gamma \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} \quad (7.1)$$

where S_{11} and S_{21} are respectively the reflection and transmission coefficients for light incident on the medium from side 1; S_{22} and S_{12} being the same for incidence on side 2. In the absence of any magneto-optical effects, reciprocity dictates that $S_{12} = S_{21}$, so only three independent parameters are necessary to describe the optical properties of such a device.

An effective nonlinearity is derived, in this simple device geometry, from the coherent nature of beam interactions and the properties of matrix (1): The linearity of the film's properties and thereby the scattering matrix imply trivially that if both input signals are scaled by a factor η , both output signals must do the same, i.e.

$$\begin{bmatrix} \eta E_\delta \\ \eta E_\gamma \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} \eta E_\alpha \\ \eta E_\beta \end{bmatrix} \quad (7.2)$$

But there is nothing to dictate any proportional scaling of one, other or both outputs when only one of the inputs is changed. So it is generally the case that:

$$\begin{bmatrix} \eta E_\delta \\ \eta E_\gamma \end{bmatrix} \neq \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} \eta E_\alpha \\ E_\beta \end{bmatrix} \quad (7.3)$$

Thus, the relationship between a given input port and a given output port in a coherent four-port device based upon a linear medium can, somewhat counterintuitively, be nonlinear. In what follows we will illustrate how the full scope of this nonlinearity can be exploited in nanostructured, subwavelength thickness media to facilitate the optical realization of several key signal processing functions.

7.1.4 New Opportunities Presented by Nano-engineered Ultrathin Media

In principle, the effective nonlinearity of a four-port device can be observed using any interaction medium: (7.1) places no constraints on the thickness of the film, which may be a single- or multi-layered, and each layer may be homogenous or nanostructured (though only at the subwavelength scale to exclude diffraction effects); the only requirement is that the incident beams must be coherent.

However, the full extent of coherent control can be realized only in the contrast between the strength of light-matter interactions in the nodes and antinodes of standing wave illumination. The challenge therefore lies in achieving the requisite balance among reflection, transmission and absorption in a medium of sufficiently subwavelength optical thickness to harness this contrast. In the optical (VIS/NIR) domain this is more challenging than one might imagine because, in the limit of very small but nonetheless finite thickness, as Sect. 7.2.3 will illustrate, nature offers a surprisingly limited palette of properties—materials are generally either highly transparent or highly reflecting; significant levels of single-beam, single-pass absorption are very difficult to achieve.

Planar metamaterials or metasurfaces though—man-made media structured on the subwavelength scale—provide unprecedented freedom to engineer the balance among resonant absorption, reflection and transmission almost at will in ultrathin metal, dielectric and semiconductor films, thereby uniquely enabling four-port optical devices to be configured to deliver a variety of nonlinear optical signal processing function (Sect. 7.3).

7.2 Coherent Interactions on Ultrathin Films

7.2.1 Singular Scattering Coefficients for Ultrathin Media

Consider now a target film of sufficient thinness that retardation effects across it can be ignored, i.e. such that we may consider each constituent molecule to be exposed to the same electric field—the combined field $E_\alpha + E_\beta$ of the incident waves. Absorbed energy will be re-radiated equally in the forward and backward directions with an efficiency dependent on the excitation wavelength λ and proportional to the driving field, i.e. to $s(\lambda)(E_\alpha + E_\beta)$, where $s(\lambda)$ is a complex wavelength-dependent scattering coefficient for the film under single-beam illumination [36]. The magnitude and phase of $s(\lambda)$ respectively correspond to the relative amplitude of the re-radiated field and to the phase lag between re-radiated and driving fields. $s(\lambda)$ may include losses and thereby excludes any assumption of equality between the combined incident and combined output beam intensities. Requirements of field continuity then dictate the following *exact* scattering matrix expression relating input and output fields for a coherent four-port device based upon a vanishingly thin film (Fig. 7.1b):

$$\begin{bmatrix} E_\delta \\ E_\gamma \end{bmatrix} = \begin{bmatrix} s(\lambda) & s(\lambda) + 1 \\ s(\lambda) + 1 & s(\lambda) \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} \quad (7.4)$$

(As it must, for single-beam illumination—with one or other of the inputs set to zero, this expression reduces to stipulate that the reflected field is equal to the re-radiated field and the transmitted field to the sum of incident and re-radiated fields.).

In cases where the contribution from interference among multiply reflected/transmitted beams is small, (7.4) can serve as a good approximation to the analytical expression of (7.1) for realistic (i.e. finite as opposed to infinitely thin) materials. This greatly simplifies the consideration and analysis of four-port coherent device functionalities (Sect. 7.3) by reducing the number of free parameters required to describe the system from three (S_{ij}) to one, in the form of $s(\lambda)$.

At the node of a standing wave formed by counter-propagating coherent incident beams, $E_\alpha = -E_\beta$. With an ultrathin medium positioned at the node, E_γ will thus always be equal to E_α , and E_δ to E_β , regardless to the value of the film's scattering parameter $s(\lambda)$. This is the “coherent perfect transmission” regime introduced in Sect. 7.1.2 above—a situation in which there is no light-matter interaction because the film is located at a point where the net electric field is zero. At an anti-node of the standing wave, on the other hand, $E_\alpha = E_\beta$. Here, both E_γ and E_δ can be reduced to zero if $\text{Re}\{s(\lambda)\} = -0.5$ and $\text{Im}\{s(\lambda)\} = 0$, whereby the film exhibits the maximum possible level of zero-thickness single-beam absorption [37, 38], namely 50%. This is the regime of “coherent perfect absorption”. It is interesting to note here that while perfect transparency requires an absorber of substantially subwavelength thickness [11], the phenomenon coherent perfect absorption does not. Indeed,

the latter has been demonstrated in a variety of optically thick media and spectral domains [39–41]. Any ultrathin film will manifest coherent perfect transparency at a standing wave node, but via metamaterial nanostructuring one may engineer a level of single-beam absorption in a film such that it will also manifest perfect coherent absorption at an anti-node. These node/anti-node limiting cases of standing wave light-matter interactions can be understood intuitively, but in what follows we also consider intermediate values of the incident waves' mutual phase $\theta = \text{Arg}\{E_\alpha\} - \text{Arg}\{E_\beta\}$ in the plane of ultrathin target media and their relevance to nonlinear signal processing functions.

7.2.2 Effective Nonlinearity in Ultrathin Films

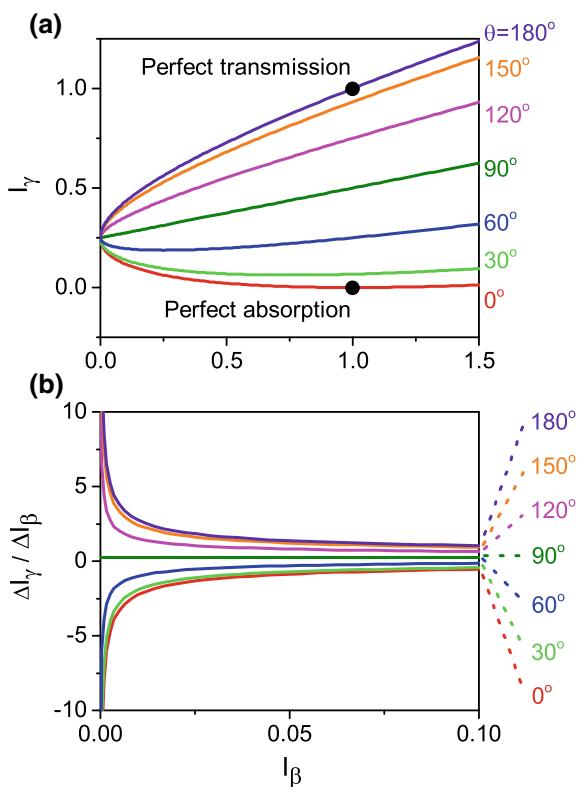
The effective optical nonlinearity of coherent four-port devices—the nonlinear relationship between selected input and output port intensities, which again should not be confused with the purely linear nature of the thin-film medium's optical response, is readily illustrated in the limit of deeply subwavelength interaction media, i.e. on the basis of (7.4) as opposed to (7.1). Consider the intensity (defined as $I = EE^*$) of output γ :

$$I_\gamma = |1 + s(\lambda)|^2 I_\alpha + |s(\lambda)|^2 I_\beta + 2\text{Re}\left\{(1 + s(\lambda))s^*(\lambda)E_\alpha E_\beta^*\right\} \quad (7.5)$$

Clearly, if one of the two input signals is removed, the system reverts to a truly linear single-beam mode of operation wherein output signal intensities are strictly proportional to that of the remaining input. And, if input intensities I_α and I_β are increased or decreased in proportion, there will be a correspondingly proportional increase or decrease in I_γ . However, if I_α remains fixed and only I_β changes, then I_γ will respond in a nonlinear fashion as illustrated in Fig. 7.2 for the case where $s(\lambda) = -0.5 + 0i$ and $I_\alpha = 1$: here the dependence of I_γ on I_β is not only nonlinear but also nonmonotonic for $\theta < 90^\circ$.

Light propagating in a conventional nonlinear (bulk) medium experiences harmonic distortion, which can lead ultimately to optical instability and multi-stability. The nonlinear character of a four-port coherent device however, is different: it is based upon the re-distribution of energy among ports and does not cause harmonic distortion; it is underpinned only by linear interference effects and as such is of course strictly compliant with energy conservation; a variety of functionalities are enabled by the fact that the level of absorption in such devices is not fixed, but is instead strongly dependent on the mutual intensity and phase of input beams.

Fig. 7.2 Nonlinear response function of a generic four-port coherent device. Dependence of output intensity I_γ on input intensity I_β (a) for a fixed input intensity $I_\alpha = 1$, and corresponding differential change curves (b), for a selection of mutual phase settings θ between E_α and E_β (as labelled) in the plane of a vanishingly thin absorber with a scattering parameter $s(\lambda) = -0.5$. [Figure 7.2a is reprinted with permission from ACS Photon. 4, 3000 (2017). Copyright of American Chemical Society]



7.2.3 Scattering Coefficients for Real (Finite) Thin Film Media

It is apparent from the above analysis that generically an effective nonlinearity may be obtained using any homogenous ultrathin film including, for example, a 50:50 beamsplitter. However, in the absence of absorption there can be no coherent modulation of absorption (the maximum achievable level of coherent absorption being twice the film's single beam absorption level). The range of achievable input-output functions is then constrained effectively to those of a conventional interferometer. In this context, photonic metamaterials are an enabling technology, providing access to an otherwise inaccessible range of ultrathin-film scattering coefficients $s(\lambda)$ at any desired wavelength by nanostructural design, and thereby to a wide range of nonlinear response functions applicable to all-optical signal processing functionalities.

Before addressing specific metamaterial designs and the performance of coherent four-port devices based upon such structures, we shall consider the meaning of $s(\lambda)$ scattering coefficient values, with reference to a selection of more familiar materials

in ultrathin film form, looking first at levels of single-beam illumination, single-pass absorption given by the expression

$$A = 1 - |s(\lambda)|^2 - |s(\lambda) + 1|^2 \quad (7.6)$$

It should be emphasised that $s(\lambda)$ is purely a material parameter, it assumes nothing as to the mode of illumination (single-beam or coherent). Figure 7.3 however maps levels of single-beam absorption onto the complex $s(\lambda)$ parameter space.

The circular line in Fig. 7.3 satisfies the condition $A = 0$, which is to say that it describes lossless media that neither absorb nor amplify incident light. It defines a boundary between absorbing materials ($A > 0$, inside the circle) and gain media ($A < 0$, outside the circle—not considered further in this work). At the central point of the circle $A = 0.5$, i.e. the maximum possible level of single-beam absorption permitted by field continuity constraints in a vanishingly thin film [37, 38]. This point corresponds to a material film with a scattering parameter $s(\lambda) = -0.5$.

A thin film of empty space (i.e. vacuum) will obviously transmit 100% of incident light at any wavelength without changing its phase. It has a value $|s(\lambda)| = 0$ it is represented in the $s(\lambda)$ plane by a single point on the zero-loss contour as shown in Figs. 7.3 and 7.4. A thin film of perfect electric conductor (PEC), on the other hand, will reflect 100% of incident light at any wavelength with a π phase change. With a value $|s(\lambda)| = -1$ it is represented by a point diametrically opposite vacuum on the zero-loss line in $s(\lambda)$ space. Thin, ideally transparent glass films with a fixed refractive index of 1.5 also sit on the zero-loss contour, being represented over a given wavelength range (for a given thickness) by a line as opposed to a singular point—traces are plotted in Fig. 7.4 for glass films of 10 and 50 nm thicknesses for wavelengths from 750 to 1050 nm. Gold thin films are again dispersive but highly reflective and weakly absorbing, so they appear in $s(\lambda)$ space as lines lying just inside

Fig. 7.3 Single-beam absorption in $s(\lambda)$ space. Level of single-beam illumination, single-pass absorption for ultrathin films mapped onto the complex plane of $s(\lambda)$ scattering coefficient. [Reprinted with permission from ACS Photon. 4, 3000 (2017). Copyright of American Chemical Society]

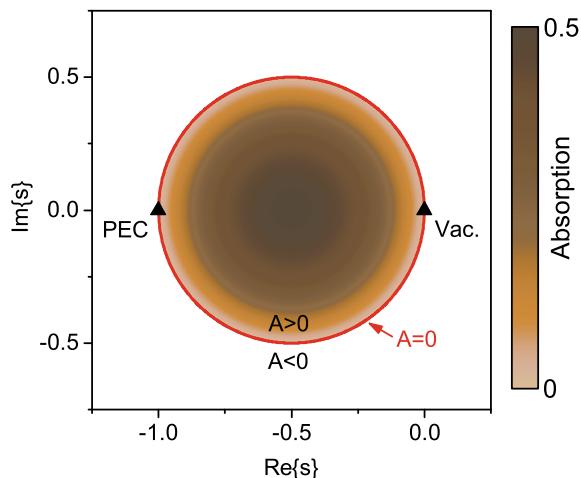
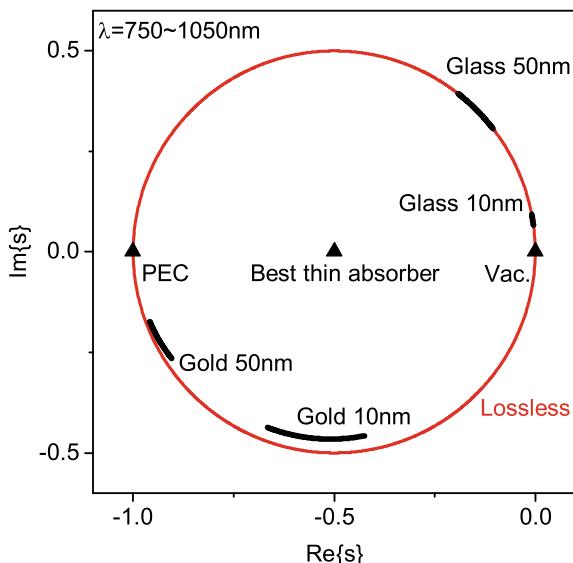


Fig. 7.4 Scattering coefficients for homogenous thin film media. $s(\lambda)$ parameters plotted for a selection of materials (as labelled) in the 750–1050 nm wavelength range. [Reprinted with permission from ACS Photon. 4, 3000 (2017). Copyright of American Chemical Society]



the zero-loss contour. Indeed, in the zero-thickness limit, most naturally occurring homogenous materials have weakly dispersive $s(\lambda)$ coefficients close to the zero-loss contour. In contrast, metamaterial nanostructures can provide access by design to the full parameter space.

7.2.4 Scattering Coefficients for Nano-engineered Metasurfaces

It is apparent that in the limit of substantially subwavelength, but nonetheless finite, thickness the $s(\lambda)$ approximation to the full scattering matrix description of materials holds for homogenous, weakly dispersive media. But a question arises as to whether it remains valid for resonant, highly dispersive, finite thickness metamaterials.

We first consider first a generic plasmonic metamaterial design, comprising an array of asymmetric split ring (ASR) resonator slits cut through a thin noble metal film (50 nm gold) on a thin transparent dielectric (30 nm silicon nitride) substrate as illustrated in Fig. 7.5a, which has been employed and optimized in numerous previous studies [31, 42], including the first experimental demonstrations of coherent absorption modulation [11, 12].

The single-beam reflection, transmission and absorption characteristics of this metamaterial type are well-known and can be numerically simulated, as shown in

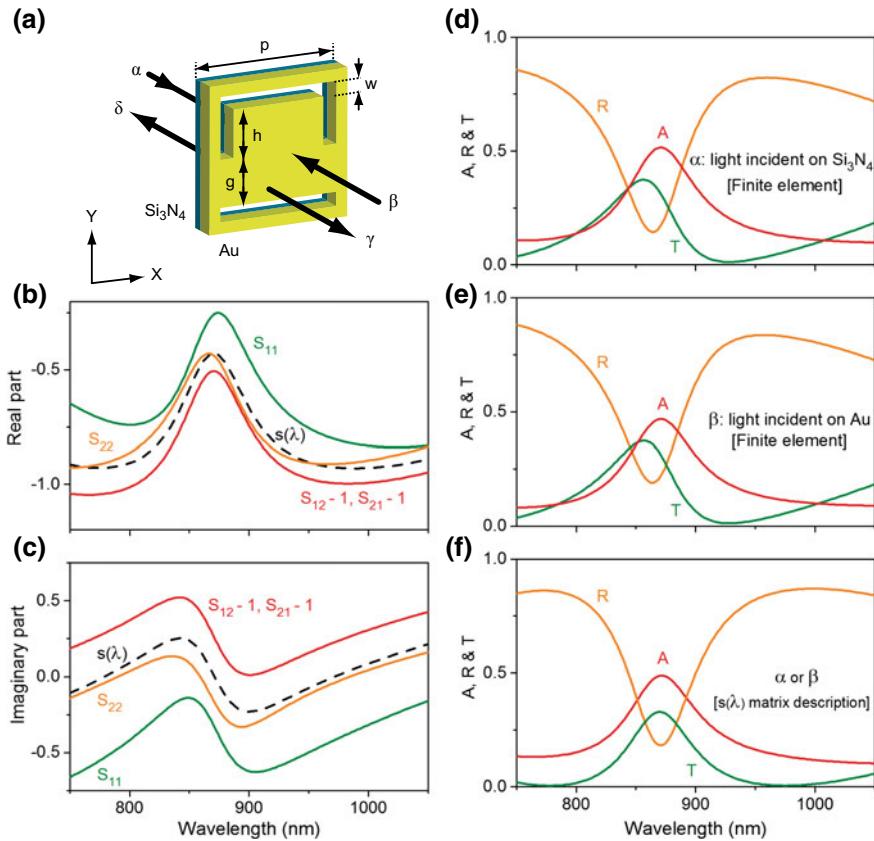


Fig. 7.5 The $s(\lambda)$ approximation applied to a realistic photonic metamaterial. **a** Gold-on-silicon nitride asymmetric split ring metamaterial unit cell geometry: cell size $p = 430$ nm; linewidth $w = 40$ nm; arm and gap sizes $h, g = 140$ nm. **b, c** Spectral dispersion of the real and imaginary parts of the metamaterial's S_{ij} scattering coefficients (as labelled), and that of the corresponding $s_M(\lambda)$ scattering parameter evaluated according to (7.7). **d, e** Numerically simulated single-beam reflection R , transmission T and absorption A spectra for waves α and β incident respectively on the substrate (silicon nitride) and metal film sides of the metamaterial array. **f** Propagation direction-independent R , T and A spectra derived using the simplified $s_M(\lambda)$ scattering parameter description of the metamaterial

Fig. 7.5d, e. This model, implemented using the finite element method (in COMSOL Multiphysics), imposes periodic boundary conditions in the x and y directions as defined in the unit cell schematic of Fig. 7.5a (i.e. assumes an infinite planar metamaterial array), employs a Drude-Lorentz model for the properties of gold and a fixed refractive index of 2.0 for silicon nitride, and assumes monochromatic plane-wave, y -polarized illumination along the surface-normal ($\pm z$) direction.

ASR metamaterials characteristically present a well-defined Fano-type absorption resonance at a wavelength set by the material composition and geometry (dimensions) of the unit cell. The finite thickness of the model system and its directional asymmetry (the fact that it is a bilayer construction with gold on one side and silicon nitride on the other) are manifested in the fact that the reflection and absorption coefficients for the two propagation directions of light differ slightly (but as expected in a linear reciprocal system the transmission coefficients are identical).

Nevertheless, the properties of the metamaterial can reasonably be encapsulated in a simplified wavelength-dependent complex scattering parameter $s_M(\lambda)$. The spectral dispersion of the metamaterial's full S_{ij} scattering coefficients is presented in Fig. 7.5b, c, overlaid with that of $s_M(\lambda)$, which we define on the basis of an equally weighted average over the S_{ij} coefficients:

$$s_M(\lambda) = \frac{1}{4} [S_{11} + (S_{12} - 1) + (S_{21} - 1) + S_{22}] = \frac{1}{4} [S_{11} + S_{12} + S_{21} + S_{22} - 2] \quad (7.7)$$

Directionally-independent reflection, transmission and absorption spectra for the metamaterial, calculated using $s_M(\lambda)$, are presented in Fig. 7.5f. Their close agreement with the numerically modelled spectra of Fig. 7.5d, e demonstrate that the scattering matrix of (7.4) and associated scattering parameter $s_M(\lambda)$, can reasonably approximate the characteristics of realistic plasmonic metamaterial thin films.

As discussed in previous sections, the great power of the metamaterial paradigm lies in its ability to provide otherwise inaccessible optical properties by nanostructural design. This is illustrated by the trace in Fig. 7.6 corresponding to the above ASR metamaterial in the $s(\lambda)$ plane—for wavelengths between 750 and 1050 nm it describes a loop extending from the neighbourhood of the zero-loss contour, towards the centre of the $A > 0$ domain and ‘ideal, zero-thickness absorber’ characteristics, and back. This trace is $s(\lambda)$ space can be modified simply by adjusting the dimensions of the metamaterial structure: the ASR2 line for example represents a gold/silicon nitride metamaterial design with an absorption resonance centred at 1550 nm. And of course, the ASR is only one of an almost infinite number of possible designs—a third trace is presented in Fig. 7.6 for a metamaterial comprising an array of linear slots again in a gold film on silicon nitride.

7.2.5 Scattering Coefficients for Lossless Media

It was established at the outset, in Sect. 7.2.1, that the $s(\lambda)$ scattering parameter incorporates losses, and it has been seen that the matrix

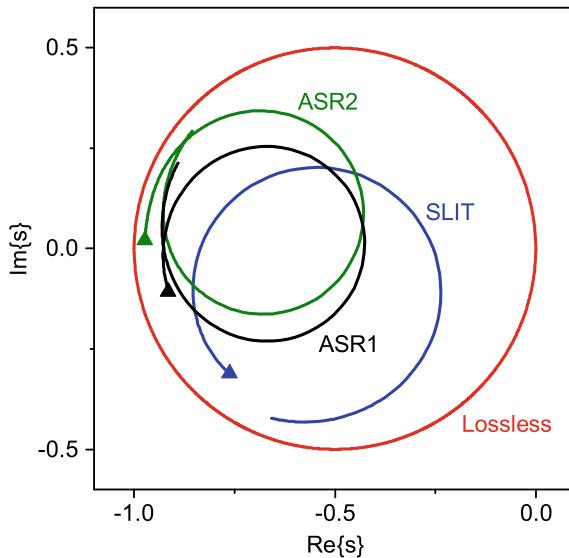


Fig. 7.6 Scattering coefficients for planar photonic metamaterials. Dispersion curves for the $s_M(\lambda)$ scattering parameters of the ASR metamaterial design of Fig. 7.5 (curve ASR1), another 50 nm gold/30 nm silicon nitride ASR metamaterial design scaled ($p = 1000$ nm; $w = 100$ nm; $h, g = 350$ nm) to provide an absorption resonance at 1550 nm (ASR2, plotted for the wavelength range 1250–1850 nm), and a metamaterial array of nano-slits (in 30 nm gold on 30 nm silicon nitride; period = 430 nm; slot length $x/\text{width } y = 200/60$ nm; plotted for 750–1050 nm). [Triangular markers denote the short-wavelength end of each dispersion curve; N.B. wavelength is not directly proportional to distance along the line.] [Figure 7.6a is reprinted with permission from ACS Photon. 4, 3000 (2017). Copyright of American Chemical Society]

$$\begin{bmatrix} s(\lambda) & s(\lambda) + 1 \\ s(\lambda) + 1 & s(\lambda) \end{bmatrix} \quad (7.8)$$

provides a good approximation to the full scattering matrix

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \quad (7.9)$$

description of the properties of real, finite thickness, homogenous and metamaterial nanostructured media.

Interestingly, in the specific case of zero-loss ultrathin media there is a more direct correlation. Energy conservation and reciprocity considerations [43] dictate that the conventional scattering matrix of a lossless film (regardless of its thickness) must be unitary and symmetric. Moreover, if the film is symmetric in relation to the $\pm z$ propagation directions, this matrix reduces to

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{12} & S_{11} \end{bmatrix} \quad (7.10)$$

with components satisfying the conditions

$$S_{11}S_{11}^* + S_{12}S_{12}^* = 1 \quad (7.11)$$

$$S_{11}S_{12}^* + S_{12}S_{11}^* = 0 \quad (7.12)$$

A vanishingly thin medium is inherently propagation-direction symmetric and if lossless must have a scattering parameter $s(\lambda)$ satisfying, by analogy with (7.11) and (7.12), the conditions

$$s(\lambda)s^*(\lambda) + (1 + s(\lambda))(1 + s^*(\lambda)) = 1 \quad (7.13)$$

$$s(\lambda)(1 + s^*(\lambda)) + s^*(\lambda)(1 + s(\lambda)) = 0 \quad (7.14)$$

which, as a validation of the $s(\lambda)$ analytical framework, reduce to the same expression defining the $A = 0$ contour in the $s(\lambda)$ parameter map.

7.2.6 Magnetic Field Scattering Coefficients

We have so far, and will in subsequent sections, restricted ourselves to media with an electric dipolar optical response, but it is interesting to consider what would happen were the dominant response of the medium of a magnetic nature, such as in some metamaterials [44, 45] or at magnetic-dipolar resonances in solids [46, 47].

The analysis is analogous to that for electric dipolar media, producing a matrix expression of familiar form relating the input and output magnetic fields of the four-port device:

$$\begin{bmatrix} B_\delta \\ B_\gamma \end{bmatrix} = \begin{bmatrix} m(\lambda) & m(\lambda) + 1 \\ m(\lambda) + 1 & m(\lambda) \end{bmatrix} \begin{bmatrix} B_\alpha \\ B_\beta \end{bmatrix} \quad (7.15)$$

Here $m(\lambda)$ is the corresponding wavelength-dependent complex magnetic scattering parameter. In terms of electric fields, accounting for the mutual orientation of the electric and magnetic field components, this becomes

$$\begin{bmatrix} E_\delta \\ E_\gamma \end{bmatrix} = \begin{bmatrix} -m(\lambda) & m(\lambda) + 1 \\ m(\lambda) + 1 & -m(\lambda) \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} \quad (7.16)$$

Note the reversal of the sign of the diagonal terms. This has an important implication in that it dictates that the maximum of coherent absorption for a

magnetic-dipolar thin film will be achieved at the nodes of the electric field standing wave, as opposed to its anti-nodes where coherent absorption is maximized for electric-dipolar thin films. As for electric-dipolar media, the maximum level of single-beam (traveling wave) magnetic absorption is 50%.

The relation between inputs (α and β) and outputs (δ and γ) is more complex if the medium responds to both electric and magnetic fields. In this case, the electric fields radiated by the medium in either direction are no longer symmetric (as they are in the case of a purely electric-dipolar response) or anti-symmetric (as in the case of a purely magnetic response). Their relative magnitude and phase will depend on several properties of the driven electric and magnetic dipoles, such as their relative strengths, their phase lag in excitation, and whether they mutually couple in the near-field. For example, it has recently been shown in the microwave domain that the field radiated in the backward direction by an appropriately designed metamaterial illuminated by a single beam (traveling wave) can be zero while in the forward direction it cancels that of the original travelling wave. As such, the wave is perfectly absorbed by the medium [48].

7.2.7 Cascading Metadevices

Performing complex optical data processing functions with coherent four-port metadevices will ultimately require cascaded arrangements (sequential operation) of several individual elements (c.f. electronic logic circuits comprising numerous individual logic gates). This represents a very substantial practical challenge, not least in the need for the output from one device to be suitable as the input to the next, or indeed the next two [24]. At a more basic level, cascading metamaterial elements provides an additional means (aside from modifying metamaterial design) of accessing the full $s(\lambda)$ parameter space to obtain specific input/output nonlinear response functions. Indeed, one may envisage reconfigurable cascades wherein individual elements can be switched in or out of the chain to modify the overall response function. The $s(\lambda)$ matrix formalism provides an elegant framework of analysing cascaded devices.

Consider a cascade of n coherently illuminated metamaterial elements, such as in the waveguide network implementation illustrated schematically in Fig. 7.7a, where circulators provide for the requisite routing of signals from one element to the next (for simplicity we ignore the phase accumulated in propagation, effectively assuming that the metamaterials are stacked sequentially front-to-back). Based on the scattering parameters $s_1(\lambda)$, $s_2(\lambda)$, $s_3(\lambda)$, etc. for the individual elements (which may all be different), a single effective scattering parameter $s(\lambda)$ can be obtained describing the entire network—relating the inputs (α and β) to the first element to the outputs (γ and δ) from the last. For odd values of n , the cascade matrix retains the same form as the single-element matrices:

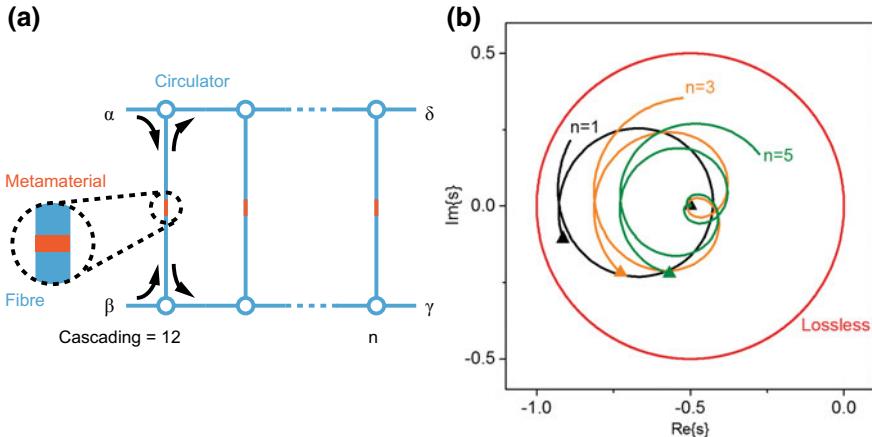


Fig. 7.7 Effective scattering parameters for cascaded metadevices. **a** Schematic network implementation of a coherent four-port metadevice cascade comprising n stages. **b** Dispersion curves for the scattering parameters for a single ASR metamaterial element of the design presented in Fig. 7.5 and for cascades of $n = 3$ and $n = 5$ such elements (as labelled), for the wavelength range from 750 nm (denoted by the triangular markers) to 1050 nm

$$\begin{bmatrix} E_\delta \\ E_\gamma \end{bmatrix} = \begin{bmatrix} s'(\lambda) & s'(\lambda) + 1 \\ s'(\lambda) + 1 & s'(\lambda) \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} = \dots \begin{bmatrix} s_3(\lambda) & s_3(\lambda) + 1 \\ s_3(\lambda) + 1 & s_3(\lambda) \end{bmatrix} \begin{bmatrix} s_2(\lambda) & s_2(\lambda) + 1 \\ s_2(\lambda) + 1 & s_2(\lambda) \end{bmatrix} \dots \begin{bmatrix} s_1(\lambda) & s_1(\lambda) + 1 \\ s_1(\lambda) + 1 & s_1(\lambda) \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} \quad (7.17)$$

In the simplest case where the n individual metamaterial elements are identical

$$\begin{bmatrix} E_\delta \\ E_\gamma \end{bmatrix} = \begin{bmatrix} s'(\lambda) & s'(\lambda) + 1 \\ s'(\lambda) + 1 & s'(\lambda) \end{bmatrix} \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} = \begin{bmatrix} s(\lambda) & s(\lambda) + 1 \\ s(\lambda) + 1 & s(\lambda) \end{bmatrix}^n \begin{bmatrix} E_\alpha \\ E_\beta \end{bmatrix} \quad (7.18)$$

Figure 7.7b presents traces in the $s(\lambda)$ parameter space for cascades of the ASR metamaterial absorber of Fig. 7.5 above, again for the wavelength range from 750–1050 nm, alongside the curve for a single element, illustrating how the cascades can provide access to a wide range of $s(\lambda)$ values on the basis of a single metamaterial design.

7.3 Coherent Metadevices

7.3.1 Small Signal Amplification

On the basis of the nanostructurally engineered nonlinear input-output relationships achievable in ultrathin-film coherent four-port devices a variety of all-optical signal switching and processing functions may be realized, mimicking the functionalities of electronic circuit elements, as shown in [49]: For example, a coherent optical metadevice may replicate the small-signal amplification function of a transistor, whereby small intensity modulations at a given input translate to large amplitude modulations at a given output, if for a fixed (bias) input intensity I_α , the output signal intensity change ΔI_γ is greater than the input signal change ΔI_β . (Because the output level of the four-port optical device is biased it cannot be employed as a DC amplifier.)

From (7.5), the small-signal differential gain $G = \Delta I_\gamma / \Delta I_\beta$ is

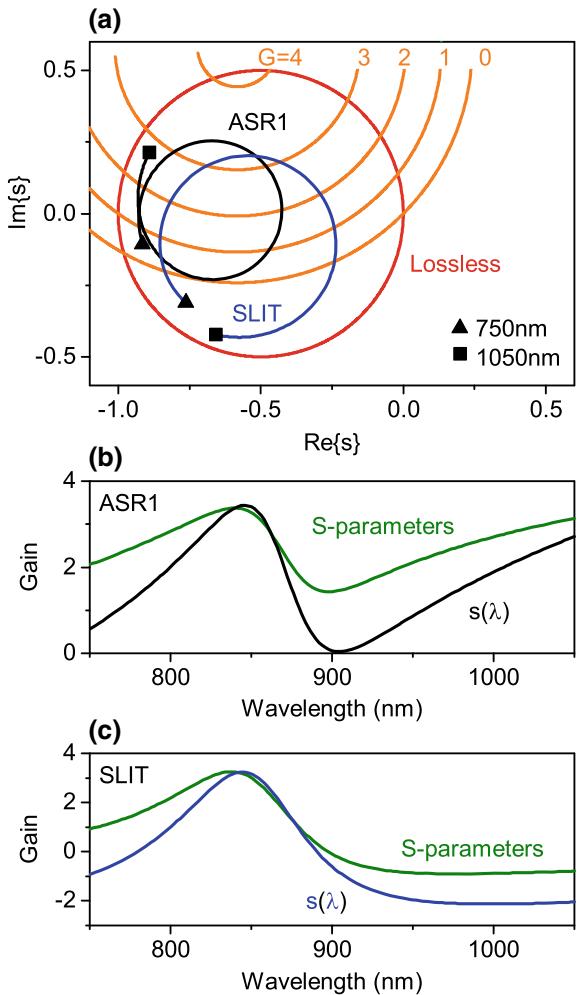
$$G = \frac{dI_\gamma}{dI_\beta} = |s(\lambda)|^2 + \operatorname{Re} \left\{ (1 + s(\lambda))s(\lambda)^* e^{i\theta} \right\} \sqrt{\frac{I_\alpha}{I_\beta}} \quad (7.19)$$

G can in principle become infinitely large as I_β tends to zero. The level of optical gain that can be achieved using selected metasurface interaction media (the ASR1 and SLIT designs introduced in Sect. 7.2.4) are illustrated in Fig. 7.8a, where gain contours from (7.19) are plotted in the complex $s(\lambda)$ plane for a specific input intensity ratio $I_\alpha/I_\beta = 100$ and setting of the mutual phase $\theta = 3\pi/4$ between incident waves at the metamaterial plane.

Under these circumstances the differential gain for the ASR1 metamaterial design reaches a maximum value, calculated either on the basis of the single $s_M(\lambda)$ scattering parameter of (7.7) or the full set of S_{ij} coefficients for the metamaterial (i.e. taking into account the directional asymmetry of the gold/silicon nitride bilayer illustrated in Fig. 7.5d, e), of ~ 3.4 at a wavelength λ of ~ 840 nm as shown in Fig. 7.8b.

In the manner of Figs. 7.2, 7.9a presents a family of nonlinear output intensity I_γ versus input intensity I_β curves (assuming a fixed input intensity $I_\alpha = 1$) for a selection of phase retardations θ between the input fields E_α and E_β at the metamaterial plane, for the $\lambda = 846$ nm peak-gain wavelength given by the $s_M(\lambda)$ scattering parameter. It is seen here that for an input retardation $\theta = 3\pi/4$ (135° , as set in Fig. 12.8) gain, i.e. values of $G > 1$, will be observed for signal input intensities below $I_\beta = 1.3$ (ratios $I_\alpha/I_\beta > i\theta/3$). Importantly, for the selected input retardation $\theta = 135^\circ$, the phase response of the amplifier (Fig. 7.9b), that is the phase difference φ between the input and output signals (β and γ), is a near-flat function of I_β , with at a value $\varphi \sim 165^\circ$, meaning that the amplified output signal will be consistently in near-antiphase with the input.

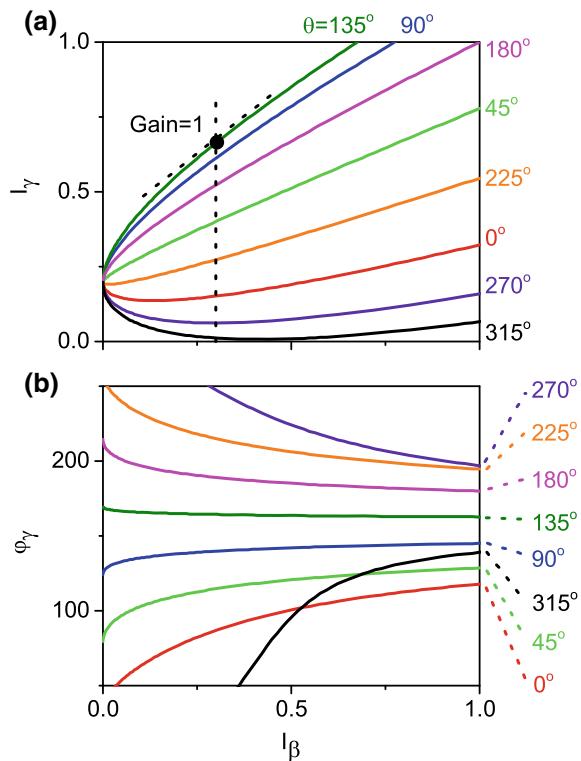
Fig. 7.8 Small-signal amplification in a four-port coherent optical device. **a** Integer differential gain contours for $G = 0 - 4$ from (7.19), for an input intensity ratio $I_\alpha/I_\beta = 100$ and mutual phase $\theta = 3\pi/4$, in the $s(\lambda)$ plane overlaid with scattering parameter dispersion traces for the ASR1 and SLIT metasurfaces first presented in Fig. 7.6. **b, c** Dispersion of G for the metasurfaces derived from their approximate $s_M(\lambda)$ [black line in (b) for ASR1; blue in (c) for SLIT] and from their full S_{ij} coefficients [green] (Color figure online)



7.3.2 Logic Operations

The nonlinear response of coherent optical four-port metadevices can also be engaged to reproduce the functionality of logic gates, as has recently been demonstrated in an optical fibre network operating in the near-IR telecoms C-band at 40 Gbit/s [50]. An AND gate—configured to produce an output signal at port γ only when both inputs (α and β) are present, according the following binary truth table:

Fig. 7.9 Four-port coherent optical amplifier performance. **a** Output intensity I_γ and **b** output phase ϕ_γ , as functions of input intensity I_β (for fixed $I_\alpha = 1$) at the $s_M(\lambda)$ gain peak wavelength of 846 nm [see Fig. 7.8]. Curves are presented in each case for a selection of different input phase retardations θ between E_α and E_β [as labelled]



Input α	Input β	Output γ
1	1	1
1	0	0
0	1	0
0	0	0

These binary logic states can be defined optically in terms of intensities $I = EE^*$: the inputs $E_\alpha E_\alpha^*$ and $E_\beta E_\beta^*$ each set to zero or I_{in} (input logic states “0” and “1” respectively); the output state of the metadevice coherent gate then being considered a logical “0” if $I_{\text{out}} = E_\gamma E_\gamma^* < \mu I_{\text{in}}$, where μ is a discrimination level that may be set between 0 and 2, and a logical “1” if the output intensity $I_{\text{out}} > \mu I_{\text{in}}$.

Following from the truth table, (16.4) and the necessity of compliance with energy conservation, a set of simultaneous requirements on the scattering parameter $s(\lambda)$ of the interaction medium needed to deliver AND functionality can be defined:

$$P: \quad |1 + s(\lambda)|^2 + |s(\lambda)|^2 \leq 1 \quad (7.20)$$

$$\text{SUM1: } |(1+s(\lambda))e^{i\theta} + s(\lambda)|^2 > \mu \quad (7.21)$$

$$\text{SUM2: } |s(\lambda)|^2 < \mu \quad (7.22)$$

$$\text{SUM3: } |1 + s(\lambda)|^2 < \mu \quad (7.23)$$

Equation (7.20)—condition P—specifies that (in the absence of gain) the output power may not exceed the total input power. In other words, in the complex $s(\lambda)i$ plane, physical solutions are found inside or on the zero-loss contour. Equations (7.21–7.23)—conditions SUM1, SUM2 and SUM3—then define a set of circular domains in $s(\lambda)$ space, with diameters proportional to μ , that together outline a parameter space within which a thin film can deliver AND logic functionality: suitable material parameters are located inside the SUM2 and SUM3 contours and outside the SUM1 contour as illustrated in Fig. 7.10 for a discrimination level $\mu = 0.5$ and input phase retardation $\theta = \pi/2$.

Under these conditions, the $s_M(\lambda)$ trace of the metasurface ASR1 introduced above lies within the domain of suitable material parameters for wavelengths between 844 and 872 nm—operational band limits that can be modified by scaling or changing the metamaterial design, cascading metasurface elements (as illustrated in Sect. 7.2.7), changing the logic discrimination level μ , or the input phase.

A similar analytical procedure may be applied to the signal inversion function of a NOT gate, configured with one signal input port (α) and one signal output (γ), and the truth table:

Input α	Output γ
1	0
0	1

Once more, the input signal intensity $E_\alpha E_\alpha^*$ may be set to zero or I_{in} (logical “0” or “1” respectively), and output logic levels defined in terms of output intensity $I_{\text{out}} = E_\gamma E_\gamma^*$ and a discrimination level μ . (In the four-port coherent optical implementation a second ‘bias’ input (β) fixed at intensity I_{in} is also required.) The requirements on $s(\lambda)$ are then:

$$P: |1 + s(\lambda)|^2 + |s(\lambda)|^2 \leq 1 \quad (7.24)$$

$$\text{INV1: } |(1+s(\lambda))e^{i\theta} + s(\lambda)|^2 < \mu \quad (7.25)$$

$$\text{INV2: } |s(\lambda)|^2 > \mu \quad (7.26)$$

In this case, as illustrated in Fig. 7.11 (assuming $\mu = 0.5$ and now $\theta = \pi/4$), the necessary material parameters are provided by the ASR1 metasurface between 900 and 936 nm. Again these operational band limits are set by the metamaterial design, logic discrimination level, and the input phase θ . Indeed, with optimization of the

Fig. 7.10 AND logical gating function in a four-port coherent optical device. **a** Inclusive [P, SUM2, SUM3] and exclusive [SUM1] conditions for the delivery of AND gating functionality [assuming input phase retardation $\theta = \pi/2$ and logic discrimination level $\mu = 0.5$] mapped onto the complex $s(\lambda)$ plane and overlaid with the scattering parameter dispersion trace for the ASR1 metasurface. The hatched area represents the viable material parameter space. **b, c** Spectral dispersion of the SUM1, 2 and 3 conditions on signal output intensity I_{out} derived from the metasurface's $s_M(\lambda)$ approximation (**b**) and from its full S_{ij} coefficients (**c**), wherein viable operating ranges are defined by the discrimination level μ being below SUM1 and above SUM2 and SUM3

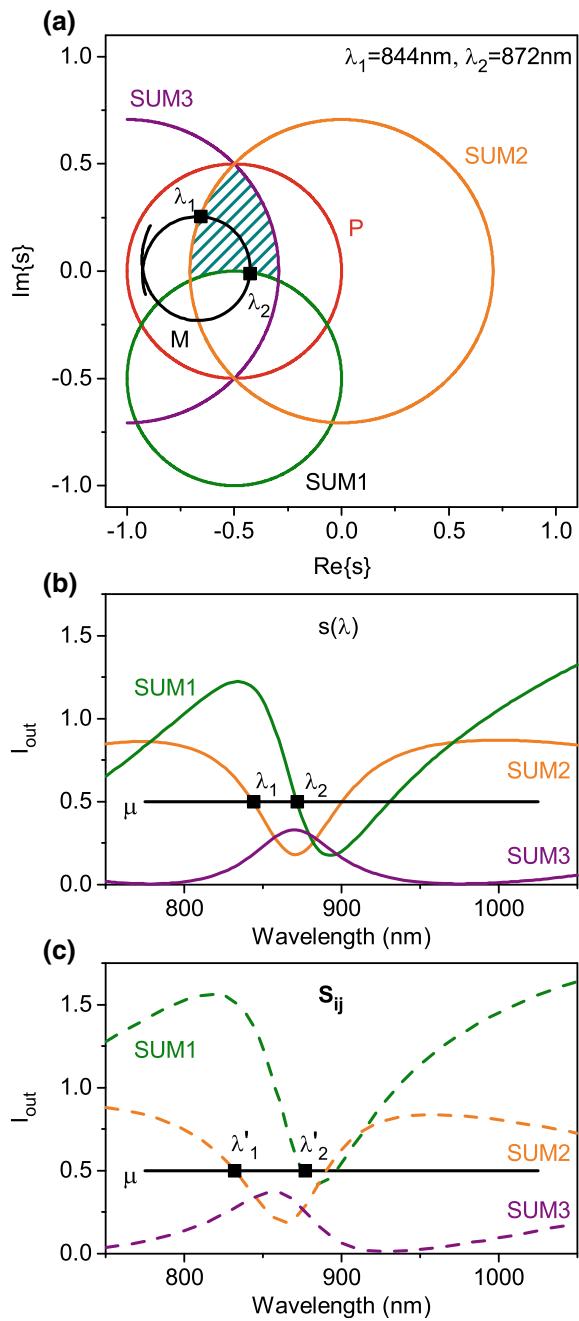
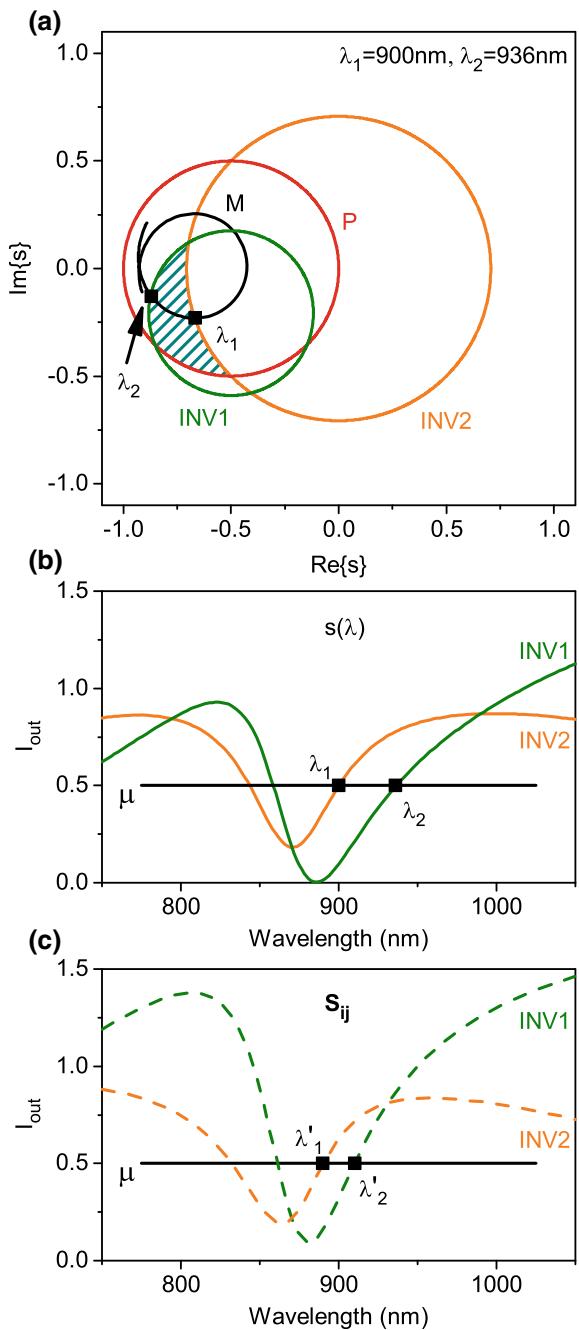


Fig. 7.11 NOT logical gating function in a four-port coherent optical device. **a** Inclusive [P, INV1] and exclusive [INV2] conditions for the delivery of NOT gating functionality [assuming input phase retardation $\theta = \pi/4$ and logic discrimination level $\mu = 0.5$] mapped onto the complex $s(\lambda)$ plane and overlaid with the scattering parameter dispersion trace for the ASR1 metasurface. The hatched area represents the viable material parameter space. **b, c** Spectral dispersion of the INV1 and 2 conditions on signal output intensity I_{out} derived from the metasurface's $s_M(\lambda)$ approximation (**b**) and from its full S_{ij} coefficients (**c**), wherein viable operating ranges are defined by the discrimination level μ being below SUM1 and above SUM2 and SUM3



metamaterial design and/or dynamic tuning of the resonant properties (e.g. through nano-mechanical spatial reconfiguration of unit cells [33, 34, 51, 52] or in hybrid systems incorporating electro-/photo-active media [53–60]) one may realize metadevices that can perform several different optical signal processing functions in a single wavelength band given appropriate settings of input signal phase, input intensity ratio and output logic discrimination level.

7.4 Coherent Perfect Absorption in the Quantum Regime

Recent experiments have shown that coherent perfect absorption, the simplest manifestation of coherent control on thin lossy films, can be observed when a single heralded photon launched into the interferometer via a lossless 50/50 beam splitter generates a coherent superposition state at a metamaterial film inserted into the interferometer [14]. This result elegantly demonstrates that while the absorption of photons from a travelling wave is probabilistic, standing wave absorption can be observed deterministically, with nearly unitary probability of coupling a photon into a mode of the material, for example a localized plasmon when this is a metamaterial excited at the plasmon resonance.

This finding also exposes the underlying quantum mechanism of optical gating via the coherent absorption process with both continuous wave and pulsed signals at classical light levels, as discussed in preceding sections. The fact that coherent absorption and thus modulation of light with light can be demonstrated with a single photon proves that the effect of modulation here does not rely on one photon modulating another, for example via an intrinsic nonlinearity of the film. Rather, coherent absorption exploits a difference in absorption probabilities between the two configurations of the gate—when the control beam is present or absent. At higher photon fluxes (i.e. at classical light levels), this takes the form of an interference-controlled redistribution of the energy flow among the inputs, the outputs and the dissipative channel provided by the absorber.

This work has recently been extended to the use of pairs of entangled photons: it has been shown that non-local polarization-sensitive detection of one of the photons in the pair predetermines the polarization state of the other, which can be used to switch ‘on’ and ‘off’ its absorption in a polarization-sensitive interferometer containing a plasmonic absorbing film [60].

7.5 Summary

We have shown here that four-port optical devices exploiting the coherent interaction of beams on ultrathin (substantially subwavelength) material films displaying purely linear optical properties (i.e. assuming no intrinsic nonlinearity or gain) can provide a strongly nonlinear input/output response function. This response function

can be controlled by design in planar metasurfaces enabling devices to be configured to deliver a variety of all-optical signal switching/processing functions analogous to the electronic functions of transistors (small signal amplification) and binary logic gates (AND and NOT functions).

Indeed, it is seen that a single generic asymmetric split ring metamaterial design can satisfy the requirements of all three of these functions in certain wavelength bands. The freedom with which the resonant dispersion of reflection, transmission and absorption can be manipulated by design in a metamaterial (ASR or otherwise), far beyond the range of properties available in homogenous naturally occurring media in the limit of deeply subwavelength thickness, enables these operational ranges to be set at any desired wavelength. Moreover, using active and reconfigurable metamaterials [33, 51–60], these characteristics, and thereby the function of four-port optical devices, may be dynamically tuned or switched.

Coherent optical gates offer advantages over nonlinearity-based signal processing devices in their freedom from harmonic distortion, their THz bandwidth, and their ability to operate down to the single-photon level. But these advantages have a price: such devices depend on the interference of light waves and as such require a coherent information carrier to be maintained across a data processing network, as well as precise positional settings to maintain relative phases. That said, locally coherent networks are already part of the photonic telecommunications roadmap where they offer increased bandwidth by providing access to additional degrees of freedom including the phase and polarization of light and to a variety of spectrally efficient modulation formats; And the issue of positional stability among optical circuit elements may be addressed in a monolithic (e.g. silicon-photonic) waveguide platform (there is nothing of fundamental importance in the free-standing nature of the metasurfaces in the proof-of-principle analyses presented here).

Four-port optical metadevices may not be best-suited to direct replication of complex electronic multi-component Boolean logic operations. Indeed, in common with any all-optical logic platform, they face implementation challenges including cascadability and input/output isolation [24]. However, they may facilitate different and novel approaches, such as via non-Boolean ‘direct logic’ data processing architectures [61] and ‘cognitive’ networks [60, 62].

References

1. M. Shapiro, P. Brumer, *Principles of the Quantum Control of Molecular Processes* (Wiley Inc., Hoboken, New Jersey, USA, 2003)
2. D.E. Reiter, E.Y. Sherman, A. Najmaie et al., Coherent control of electron propagation and capture in semiconductor heterostructures. *Europhys. Lett.* **88**, 67005 (2009)
3. A. Assion, T. Baumert, M. Bergt et al., Control of chemical reactions by feedback-optimized phase-shaped femtosecond laser pulses. *Science* **282**, 919–922 (1998)
4. M.I. Stockman, S.V. Faleev, D.J. Bergman, Coherent control of femtosecond energy localization in nanosystems. *Phys. Rev. Lett.* **88**, 067402 (2002)

5. T.S. Kao, S.D. Jenkins, J. Ruostekoski et al., Coherent control of nanoscale light localization in metamaterial: creating and positioning isolated subwavelength energy hot spots. *Phys. Rev. Lett.* **106**, 085501 (2011)
6. D. Brinks, M. Castro-Lopez, R. Hildner et al., Plasmonic antennas as design elements for coherent ultrafast nanophotonics. *Proc. Natl. Acad. Sci. USA* **110**, 18386–18390 (2013)
7. B. Gjonaj, J. Aulbach, P.M. Johnson et al., Active spatial control of plasmonic fields. *Nat. Photon.* **5**, 360–363 (2011)
8. M. Miyata, J. Takahara, Excitation control of long-range surface plasmons by two incident beams. *Opt. Express* **20**, 9493–9500 (2012)
9. S.B. Choi, D.J. Park, Y.K. Jeong et al., Directional control of surface plasmon polariton waves propagating through an asymmetric Bragg resonator. *Appl. Phys. Lett.* **94**, 063115 (2009)
10. Z. Li, S. Zhang, N.J. Halas et al., Coherent modulation of propagating plasmons in silver-nanowire-based structures. *Small* **7**, 593–596 (2011)
11. J. Zhang, K.F. MacDonald, N.I. Zheludev, Controlling light-with-light without nonlinearity. *Light. Sci. Appl.* **1**, e18 (2012)
12. X. Fang, M.L. Tseng, J.Y. Ou et al., Ultrafast all-optical switching via coherent modulation of metamaterial absorption. *Appl. Phys. Lett.* **104**, 141102 (2014)
13. A. Karvounis, V. Nalla, K. F. MacDonald, N. I. Zheludev, Ultrafast coherent absorption in diamond metamaterials. *Adv. Mater.* **30**, 1707354 (2018)
14. T. Roger, S. Vezzoli, E. Bolduc et al., Coherent perfect absorption in deeply subwavelength films in the single-photon regime. *Nat. Commun.* **6**, 7031 (2015)
15. J. Shi, X. Fang, E.T.F. Rogers et al., Coherent control of Snell's law at metasurfaces. *Opt. Express* **22**, 21051–21060 (2014)
16. S.A. Mousavi, E. Plum, J. Shi et al., Coherent control of birefringence and optical activity. *Appl. Phys. Lett.* **105**, 011906 (2014)
17. S.M. Rao, A. Lyons, T. Roger et al., Geometries for the coherent control of four-wave mixing in graphene multilayers. *Sci. Rep.* **5**, 15399 (2015)
18. M. Papaioannou, E. Plum, E. T. F. Rogers, N. I. Zheludev, All-optical dynamic focusing of light via coherent absorption in a plasmonic metasurface. *Light Sci. Appl.* **7**, 17157 (2018)
19. M. Papaioannou, E. Plum, J. Valente, et al., Two-dimensional control of light with light on metasurfaces. *Light Sci. Appl.* **5**, e16070 (2016)
20. M. Papaioannou, E. Plum, N. I. Zheludev, All-optical pattern recognition and image processing on a metamaterial beam splitter. *ACS Photon* **4**, 217–222 (2017)
21. X. Fang, M.L. Tseng, D.P. Tsai et al., Coherent excitation-selective spectroscopy of multipole resonances. *Phys. Rev. Appl.* **5**, 014010 (2016)
22. X. Fang, K. F. MacDonald, E. Plum, N. I. Zheludev, Coherent control of light-matter interactions in polarization standing waves. *Sci. Rep.* **6**, 31141 (2016)
23. M.L. Tseng, X. Fang, V. Savinov, et al., Coherent selection of invisible high-order electromagnetic excitations. *Sci. Rep.* **7**, 44488 (2017)
24. D.A.B. Miller, Are optical transistors the logical next step? *Nat. Photon.* **4**, 3–5 (2010)
25. D.J. Richardson, Filling the light pipe. *Science* **330**, 327–238 (2010)
26. S.J. Vavilov, W.L. Lewshin, *Z. Phys.* **35**, 920–926 (1926)
27. D.A.B. Miller, Device requirements for optical interconnects to silicon chips. *Proc. IEEE* **97**, 1166–1185 (2009)
28. W.F. Sharfin, M. Dagenais, Femtojoule optical switching in nonlinear semiconductor laser amplifiers. *Appl. Phys. Lett.* **48**, 321–322 (1986)
29. M. Sánchez, P. Wen, M. Gross et al., Nonlinear gain in vertical-cavity semiconductor optical amplifiers. *IEEE Photon. Tech. Lett.* **15**, 507–509 (2003)
30. K. Nozaki, T. Tanabe, A. Shinya et al., Sub-femtojoule all-optical switching using a photonic-crystal nanocavity. *Nat. Photon.* **4**, 477–483 (2010)
31. N.I. Zheludev, Y.S. Kivshar, From metamaterials to metadevices. *Nat. Mater.* **11**, 917–924 (2012)

32. R.F. Waters, P.A. Hobson, K.F. MacDonald et al., Optically switchable photonic metasurfaces. *Appl. Phys. Lett.* **107**, 081102 (2015)
33. A. Karvounis, J. Ou, W. Wu et al., Nano-optomechanical nonlinear dielectric metamaterials. *Appl. Phys. Lett.* **107**, 191110 (2015)
34. J. Ou, E. Plum, J. Zhang et al., Giant nonlinearity of an optically reconfigurable plasmonic metamaterial. *Adv. Mater.* **28**, 729–733 (2016)
35. D.M. Pozar, *Microwave Engineering* (Wiley Inc., 1998)
36. J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998)
37. C. Hägglund, S.P. Apell, B. Kasemo, Maximized optical absorption in ultrathin films and its application to plasmon-based two-dimensional photovoltaics. *Nano Lett.* **10**, 3135–3141 (2010)
38. S. Thongrattanasiri, F.H.L. Koppens, F.J. García de Abajo, Complete optical absorption in periodically patterned graphene. *Phys. Rev. Lett.* **108**, 047401 (2012)
39. W. Wan, Y. Chong, L. Ge et al., Time-reversed lasing and interferometric control of absorption. *Science* **331**, 889–892 (2011)
40. S. Dutta-Gupta, O.J.F. Martin, S. Dutta Gupta, et al., Controllable coherent perfect absorption in a composite film. *Opt. Express* **20**, 1330–1336 (2012)
41. M. Pu, Q. Feng, M. Wang et al., Ultrathin broadband nearly perfect absorber with symmetrical coherent illumination. *Opt. Express* **20**, 2246–2254 (2012)
42. E. Plum, K. Tanaka, W.T. Chen et al., A combinatorial approach to metamaterials discovery. *J. Opt.* **13**, 055102 (2011)
43. R.J. Potton, Reciprocity in optics. *Rep. Prog. Phys.* **67**, 717–754 (2004)
44. W. Cai, U.K. Chettiar, H.K. Yuan et al., Metamagnetics with rainbow colors. *Opt. Express* **15**, 3333–3341 (2007)
45. N. Liu, L. Fu, S. Kaiser et al., Plasmonic building blocks for magnetic molecules in three-dimensional optical metamaterials. *Adv. Mater.* **20**, 3859–3865 (2008)
46. N. Noginova, Y. Barnakov, H. Li et al., Effect of metallic surface on electric dipole and magnetic dipole emission transitions in Eu³⁺ doped polymeric film. *Opt. Express* **17**, 10767–10772 (2009)
47. T.H. Taminiua, S. Karaveli, N.F. van Hulst, et al., Quantifying the magnetic nature of light emission. *Nat. Commun.* **3** (2012)
48. V.S. Asadchy, I.A. Faniaye, Y. Ra'di, et al., Broadband reflectionless metasheets: frequency-selective transmission and perfect absorption. *Phys. Rev. X* **5**, 031005 (2015)
49. X. Fang, K.F. MacDonald, N.I. Zheludev, Controlling light with light using coherent metadevices: all-optical transistor, summator and inverter. *Light Sci. Appl.* **4**, e292 (2015)
50. A. Xomalis, I. Demirtzioglou, E. Plum, et al., Fibre-optic metadevice for all-optical signal modulation based on coherent absorption. *Nat. Commun.* **9**, 182 (2018)
51. J.Y. Ou, E. Plum, J. Zhang et al., An electromechanically reconfigurable plasmonic metamaterial operating in the near-infrared. *Nat. Nanotech.* **8**, 252–255 (2013)
52. J. Valente, J.Y. Ou, E. Plum et al., Reconfiguring photonic metamaterials with currents and magnetic fields. *Appl. Phys. Lett.* **106**, 111905 (2015)
53. B. Gholipour, J. Zhang, K.F. MacDonald et al., All-optical, non-volatile, bi-directional phase-change meta-switch. *Adv. Mater.* **25**, 3050–3054 (2013)
54. T. Driscoll, H.-T. Kim, B.-G. Chae et al., Memory metamaterials. *Science* **325**, 1518–1521 (2009)
55. M.J. Dicken, K. Aydin, I.M. Pryce et al., Frequency tunable near-infrared metamaterials based on VO₂ phase transition. *Opt. Express* **17**, 18330–18339 (2009)
56. B. Kang, J.H. Woo, E. Choi et al., Optical switching of near infrared light transmission in metamaterial-liquid crystal cell structure. *Opt. Express* **18**, 16492–16498 (2010)
57. O. Buchnev, J.Y. Ou, M. Kaczmarek et al., Electro-optical control in a plasmonic metamaterial hybridised with a liquid-crystal cell. *Opt. Express* **21**, 1633–1638 (2013)
58. K.M. Dani, Z. Ku, P.C. Upadhyaya et al., Subpicosecond optical switching with a negative index metamaterial. *Nano Lett.* **9**, 3565–3569 (2009)

59. D.J. Cho, W. Wu, E. Ponizovskaya et al., Ultrafast modulation of optical metamaterials. *Opt. Express* **17**, 17652–17657 (2009)
60. Y.C. Jun, E. Gonzales, J.L. Reno et al., Active tuning of mid-infrared metamaterials by electrical control of carrier densities. *Opt. Express* **20**, 1903–1911 (2012)
61. J. Hardy, J. Shamir, Optics inspired logic architecture. *Opt. Express* **15**, 150–165 (2007)
62. K. Wu, C. Soci, P.P. Shum et al., Computing matrix inversion with optical networks. *Opt. Express* **22**, 295–304 (2014)
63. C. Altuzarra, S. Vezzoli, J. Valente, et al., Coherent perfect absorption in metamaterials with entangled photons. *ACS Photon.* **4**, 2124–2128 (2017)
64. K. Wu, F.J. García de Abajo, C. Soci et al., An optical fiber network oracle for NP-complete problems. *Light Sci. Appl.* **3**, e147 (2014)

Chapter 8

Nonlinear Nanoplasmonics



Alexey V. Krasavin, Pavel Ginzburg and Anatoly V. Zayats

Abstract Recent developments in integrated optics and miniaturization of optical components and devices put forward new challenges for nonlinear optics at the (sub)wavelength scales. In order to address these challenges, plasmonic modes, related to a coupled state of photons and coherent free-carrier oscillations in conductors, their nanostructures, and plasmonic metamaterials, have recently been widely used to tailor spectral and dynamic properties of the nonlinear response. Providing strong local field enhancement, plasmonic modes boost nonlinear interactions, leading to high effective nonlinear susceptibilities and offering one of the fastest nonlinear response due to the free-carriers dynamics. In this chapter, we will overview principles and various effects in nonlinear plasmonics and plasmonic metamaterials. Engineered harmonic generation and soliton formation, related to coherent nonlinear interactions in free-electron gas are discussed and a hydrodynamic model for coherent nonlinearity is introduced. The Kerr-type nonlinearities for ultrafast optical signal processing are considered in terms of electron gas excitation and relaxation dynamics in the nanostructures. The flexibility and unique features of free-electron nonlinearities in plasmonic nanostructures are important for nonlinear plasmonic applications in free-space as well as integrated and quantum nanophotonic technologies.

A. V. Krasavin (✉) · A. V. Zayats

Department of Physics and London Centre for Nanotechnology, King's College London,
Strand, London WC2R 2LS, UK
e-mail: alexey.krasavin@kcl.ac.uk

A. V. Zayats
e-mail: a.zayats@kcl.ac.uk

P. Ginzburg
School of Electrical Engineering, Tel Aviv University, Ramat Aviv, 69978 Tel Aviv, Israel
e-mail: pginzburg@post.tau.ac.il

8.1 Introduction

Nonlinear regime of interaction of light and matter became accessible with the development of lasers. Despite nonlinear effects being weak and requiring high intensities of light to be observed, nonlinear optics enables numerous important applications in modern technologies from harmonic generation, optical parametric amplification and mode-locking in ultrafast lasers to holography, self- and cross-modulation of optical signals, and optical solitons for information processing, as well as parametric down-conversion for quantum information applications. Due to the weakness of nonlinear response of conventional materials, typically only selected, *nonlinear* materials exhibit strong nonlinear optical effects at reasonable light intensities, and typically a long propagation length in a material is needed to achieve significant nonlinear interactions [1, 2].

Recent development of integrated optics and miniaturization of optical components and devices put forward new challenges for nonlinear optics at the (sub) wavelength scales, where geometrical constraints limits the applicability of conventional macroscopic nonlinear crystals and approaches which rely on phase matching or long interaction length between optical beams. At the same time, nonlinear effects are essential for optical signal processing in photonic integrated circuits.

In order to increase the efficiency of light-matter interactions in a linear domain, plasmonic modes in conductors and their nanostructures, including metal-dielectric metamaterials, have been recently widely used [3–9]. Plasmonic excitations, either propagating surface plasmon polaritons (SPPs) on extended metals surfaces or localized surface plasmons (LSPs) of metallic nanoparticles are related to a coupled state of photons and coherent free-carrier oscillations near a conductor-dielectric interface. Plasmonic excitations provide electromagnetic field confinement near the metal interface and, as the result, the local field enhancement. This has profound consequences for nonlinear optical processes which depend on the local field intensity in a superlinear manner and, thus, can be strongly enhanced near the metal-dielectric interface in the presence of plasmonic resonances [10, 11].

Nonlinear plasmonics has been developed utilizing the above described properties of plasmonic field also taking advantage of the nanoscale range of plasmonic modes enabling to achieve nonlinear response on subwavelength scales and, thus, compatible with integrated optics [12]. There are two typical approaches: (i) to use the field enhancement provided by surface plasmons to induce nonlinear response in a nonlinear dielectric near the metal interface in hybrid metal-dielectric systems or (ii) to harvest the nonlinearity of plasmonic materials themselves. In the latter case, a nonlinear response originates from the dynamics of non-equilibrium free-electrons in metals under the influence of strong illumination. Indeed, the metal nonlinear response is one of the strongest per unit interaction length and fastest, with the response time determined by the excited electron relaxation to the ground state.

Both approaches were exploited for enhancement of coherent nonlinear interactions, such harmonics generation and wave mixing, as well as Kerr-type nonlinearities for controlling light with light. The latter is the third-order nonlinear effect leading to the modification of permittivity of the material and manifested in self- or cross-phase modulation and induced transparency or absorption.

In this chapter we will overview principles and various effects in nonlinear plasmonics and plasmonic metamaterials. After a brief introduction to basic concepts of plasmonics, we introduce the effective optical nonlinearities of metal-dielectric nanostructures arising from the field enhancement effects. A hydrodynamic model of free-electron nonlinearities is then described, needed for consideration of coherent as well as Kerr nonlinearities of metals themselves. We discuss, in turn, engineering of harmonic generation, nonlinear refraction and ultrafast all-optical switching in metallic nanostructures and plasmonic metamaterials, taking into account nonlocal response as well as the so-called epsilon-near-zero regime. Finally, nonlinear surface plasmon polaritons and plasmon-solitons are briefly discussed.

8.2 Brief Summary of Concepts of Plasmonics

8.2.1 Basic Notions of Plasmonics

Optical properties of materials with high concentration of free carriers (electrons or holes), such as metals or highly doped semiconductors, and their films and nanostructures are governed by coupling of electromagnetic field to the coherent motion of free-carrier plasma near the surface. Such systems can support different types of plasmonic modes depending on the length scale of their structuring [4, 13].

Surface plasmon polariton is a propagating surface wave at the continuous metal-dielectric interfaces (Fig. 8.1a). The SPP electromagnetic field exponentially decays on both sides of the metal interface, providing subwavelength confinement near the metal surface [4, 5]. SPP is a longitudinal wave having the components of the electric field both perpendicular to the metal interface and parallel to the wavevector, thus, the polarization of the incident light which provides the electric field components in one of those directions is required for SPP excitation. Since the SPP wavevector is larger than the wavevector of photons in the adjacent dielectric medium, coupling of light from free space into SPPs requires special arrangements such as prism or diffraction grating coupler. The large wavevector also means that SPP is a slow wave accumulating energy from incoming photons and providing the field enhancement near the metal interface compared to the field of the incident light. The SPP dispersion and, therefore, the field confinement and enhancement can be modified by structuring the interface, either metal or dielectric medium, thus achieving plasmonic waveguides (structuring along the SPP propagation direction) [5, 14] or plasmonic crystals (structuring across to the SPP propagation direction)

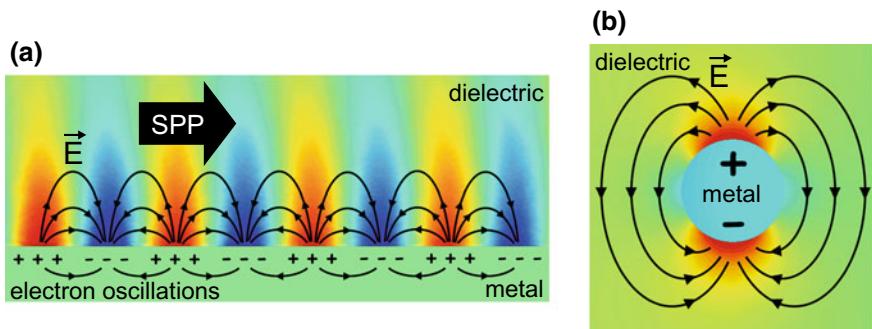


Fig. 8.1 **a** Electric field distribution (colour scale) of a surface plasmon polariton wave propagating along the interface between a metal and a dielectric and the related charge oscillations. **b** Electric field distribution of localized surface plasmon due to oscillating charges in a metal nanoparticle

[4]. In the latter case, the SPP dispersion exhibits a structure of allowed and forbidden gaps, similar to two-dimensional photonic crystals for photons.

Localized surface plasmons (LSP) are associated with the electron plasma oscillations in the confined (subwavelength) geometries, e.g., metal nanoparticles (Fig. 8.1b). LSP resonances depend on the particle size, shape and refractive index of the surroundings. LSPs can be resonantly excited with light of an appropriate frequency (and polarization, for nonspherical nanoparticles) irrespective of the excitation light wavevector. Since LSPs are confined near the nanoparticle (or a curved surface), they have small mode volume and, therefore, provide a significant electromagnetic field enhancement, which is limited by their quality factor determined by Ohmic and radiative losses as well as nonlocal effects in the case of ultrasmall sizes [15–17]. LSPs can also be assigned to the subwavelength features on a metal surface. LSP resonances play a significant role in the behaviour of SPP on rough surfaces if their frequency is close to the SPP frequency. The spectrum of LSP associated with an ensemble of metallic particles (or voids) is determined by the interaction between the individual LSP resonances. The resulting spectrum and the enhancement magnitude depend significantly on the shape and size of individual particles and a distance between them. In such ensembles of interacting small metallic particles, a very strong electromagnetic field enhancement can be observed [4, 15]. Recently, plasmonic antennas based on pairs or several plasmonic particles have been developed to control near-field to far-field light transformations [18]. A similar but simpler situation occurs in the case of a metallic particle situated above a metal surface [19]. Furthermore, sharp metal tips can produce strong local electromagnetic fields through the so-called lightning-rod effect, even when LSPs are not resonantly excited [20]. In all cases, a source of radiation localized to a sub-wavelength LSP mode can emit a signal to essentially all directions, independent of the direction of propagation of the incident field.

Plasmonic metamaterials generally consist of an ensemble of subwavelength metal nanostructures with strong interaction between surface plasmons of individual nanostructures, so that the resulting optical properties are to a great extent determined by this interaction and can be described by effective medium parameters [7]. Metamaterials based on various combinations of metallic nanostructures, such as split-ring resonators of different geometries, fishnets or nanorods, have been proposed due to the opportunities to design their optical response at will by engineering the optical response of individual components as well as electromagnetic interaction between them [7, 21].

Universal properties of all types of plasmonic nanostructures important for nonlinear optical applications are (i) the field enhancement near the metal surface compared to the free space field of the exciting light, (ii) strong sensitivity to the refractive index changes near the metal surface [22], and (iii) possibility to engineer mode dispersion and resonances by controlling nanostructure geometry and dielectric surroundings, thus, a possibility to tune the resonant response to the required operational wavelength where nonlinear response need to be enhanced. Intrinsic nonlinearity of free-electron systems is also extremely fast with the response time determined by the electron relaxation within conduction band [10], and, therefore, provides opportunities for development ultrafast all-optical effects for modulating and switching of light with light.

8.2.2 Plasmonics as a Tool for Tailoring and Enhancing Nonlinearity

Nonlinear optical interactions are relatively weak and require high light intensities to observe. The field of the incident light can however be significantly enhanced by coupling it to material structures with a resonant electromagnetic response. In particular, the electromagnetic field of plasmonic excitations is localised at the sub-wavelength scales near metal-dielectric interfaces, provides significant field enhancement and can boost nonlinear optical effects. Generally, nonlinear optical phenomena are proportional to higher powers of the driving field (e.g., a power of two for second harmonic generation (SHG) which is a second-order nonlinear process) and the induced polarization in the electric dipole approximation can be expressed as [1]

$$\mathbf{P}(\omega) = \epsilon_0 \chi^{(1)} \cdot \mathbf{E}(\omega) + \epsilon_0 \chi^{(2)} : \mathbf{E}(\omega) \mathbf{E}(\omega) + \epsilon_0 \chi^{(3)} : \mathbf{E}(\omega) \mathbf{E}(\omega) \mathbf{E}(\omega) + \dots, \quad (8.1)$$

where ϵ_0 is the permittivity of vacuum, $\mathbf{E}(\omega)$ is the electric field of illuminating light, and $\chi^{(n)}$ are the n th order optical susceptibility tensors ($\chi^{(1)}$ is the tensor of linear optical susceptibility). In order to take into account quadrupolar and magnetic dipole effects, one needs to add into (8.1) terms proportional to the field and susceptibility gradients and terms containing the magnetic field of the fundamental

wave to the respective powers (as is done for the surface-related terms in the hydrodynamic description, e.g., (8.12)).

8.2.2.1 Coherent Nonlinearities

Applying (8.1) to harmonic wave illumination $E(\omega) = E_0 \exp(i\omega t - ikx)$ with the frequency ω and the wavevector k , one can obtain $E(n\omega) \sim \chi^{(n)} E_{\text{loc}}^n(\omega)$, where $E(n\omega)$ is the field of the generated n th harmonic, and $E_{\text{loc}}(\omega)$ is the local fundamental field at the place where harmonic is generated. This local field is characterized by the frequency and position dependent enhancement factor which is related to the polarizability of the nanostructure $L(\omega, \mathbf{r}) = E_{\text{loc}}(\omega, \mathbf{r})/E_0(\omega)$, where \mathbf{r} is the position near the metal [23]. Thus, the effective nonlinear susceptibility can be significantly increased near the resonances of the plasmonic structure where the local electric field is enhanced. The generated nonlinear harmonic near-field as well as its radiation in the far-field, can further be enhanced if the electromagnetic resonances are present at the generated frequency $L(n\omega)$ [24]: $E(n\omega) \sim \chi^{(n)} L(n\omega) L^n(\omega) E_0^n(\omega)$, which can be alternatively recast in terms of the effective nonlinear susceptibility $\chi_{\text{eff}}^{(n)} = \chi^{(n)} L(n\omega) L^n(\omega)$. This motivates the quest for a local electromagnetic field enhancement achievable with various nanostructures. From this point of view, plasmonic nanostructures, which provide very high local electromagnetic fields, are perfect candidates for the realization of various concepts for augmenting of nonlinear effects [10]. It should be noted, however, that in addition to the field enhancement, the overlap of the amplitudes of the fundamental and harmonic field modes and their phase relations are important for efficient harmonic generation from nanostructures [25, 26]. The analogue of the SHG phase-matching conditions for localized plasmonic modes [27] is discussed below in Sect. 8.4.1.

8.2.2.2 Kerr-Type Nonlinearities

The Kerr-type nonlinearity can be described by the changes of the permittivity of the material under the action of control light $\mathbf{E}_c(\omega_c)$ and can be easily obtained from (8.1) keeping the leading third-order nonlinear term:

$$\mathbf{D}(\omega) = \epsilon \left(|\mathbf{E}_c(\omega_c)|^2 \right) \mathbf{E}(\omega) = \epsilon_0 \left(1 + \chi^{(1)} + \chi^{(3)} |\mathbf{E}_c(\omega_c)|^2 \right) \mathbf{E}(\omega), \quad (8.2)$$

where $\mathbf{D}(\omega)$ is the nonlinear electric displacement.

Plasmonic excitations are extremely sensitive to the refractive index changes either in the metallic structure itself or in the surrounding material [22, 28, 29]. This property can be exploited to control light with light, when a control beam induces

the nonlinear change, modifying the plasmonic resonances and through them the propagation of a signal beam in waveguides or transmission or reflection of light from plasmonic nanostructures.

Under illumination of the control light, the third-order Kerr-type nonlinearity leads to the intensity-dependent refractive index $n(I) = n_0 + \gamma |E_{\text{loc}}(\omega_c)|^2$ and absorption $\alpha(I) = \alpha_0 + \beta |E_{\text{loc}}(\omega_c)|^2$, where n_0 and α_0 are the linear refractive index and absorption, respectively, $E_{\text{loc}}(\omega_c)$ is the intensity of the control light, and γ and β are the nonlinear refraction (sometime also called n_2) and absorption coefficients. This modified (complex) refractive index can be used to alter phase or absorption of the signal light at different frequency interacting with the nanostructure (the so-called cross-modulation). The strong beam can also influence itself via the same effects, leading to the so-called self-modulation. The local fields can be related to the polarizability of the plasmonic nanostructures in the same way as discussed above, and the effective nonlinear refraction and absorption coefficients can be introduced. The field enhancements inherent to plasmonic resonances, facilitate this nonlinear effect lowering the required control light intensities.

Plasmonic field enhancement and refractive index sensitivity effects for enhancing coherent and Kerr-type nonlinearities can be applied to nonlinearities provided by the nonlinear materials next to a plasmonic nanostructure (the so-called hybrid nonlinear plasmonic systems) as well as in the case when nonlinearity of a metal itself provides a nonlinear response.

Surface enhanced Raman scattering (SERS) is one of the most famous examples of the processes enhanced by the local plasmonic field: a rough noble metal surface was shown to enhance the SERS intensity by 14 orders of magnitude compared to the conventional Raman scattering [15]. The advantage of double-resonant plasmonic structures was demonstrated for SERS, where pair of particles provides the resonant enhancement for both pump and Stokes frequencies [30]. The grooves etched in a metal surface and organized into a grating geometry have been shown to enhance the four-wave mixing (FWM) efficiency by two orders of magnitude [31]. Four orders of magnitude improvement of FWM was demonstrated with a plasmonic dimer configuration [32]. Optical nonlinearities assisted by plasmonic nanostructures have recently been proposed for various applications in active photonic components [14, 33–39], refractive-index sensing [22], chemical sensing [40] and signal processing [14, 41]. The enhanced nonlinear processes relying on the strong local field enhancement delivered by plasmonic structures have been harnessed in nonlinear materials neighbouring metal nanostructures, such as polymers [42], noble gasses [43] and others (reviewed in [10]). Nanostructuring gold in a nanorod-based plasmonic metamaterial allows one to enhance the effective third-order susceptibility of the composite media relying on free-electron nonlinearities of metal by several orders of magnitude, providing an avenue to design efficient and ultrafast nonlinear functionalities either to manipulate free-space

optical beams [44, 45] or integrated guided waves [46], as be discussed in Sect. 8.5.5. Similarly, the geometry of the unit cell in plasmonic metasurfaces can be optimized such as to simultaneously enhance both the coherent nonlinear response of these subwavelength films and their ability to produce Kerr-based ultrafast sub-picosecond all-optical modulation [47].

8.2.3 Plasmonic Metals as Nonlinear Materials

Free-electron plasma in metals and other plasmonic materials itself exhibits a strong nonlinear optical response due to complex dynamics of the electron gas in the inhomogeneous driving electromagnetic field. Free-electron nonlinearities provide one of the largest nonlinear susceptibilities and lead to a vast family of intriguing nonlinear phenomena. Furthermore, metallic nonlinearities are inherently ultrafast allowing processing of optical signals at up to a femtosecond scale, which is the second advantage of nonlinear plasmonics.

In terms of coherent nonlinearities, even-harmonic generation is forbidden in the dipolar approximation in bulk centrosymmetric materials but significant surface contributions are observed, as the surface breaks the symmetry. With the decrease of the nanostructure size, additional contributions from the bulk electron gas may arise due to inhomogeneous field and nonlocal effects. Unfortunately, even modern experimental data on the nonlinear response of metallic nanostructures from different sources are not always consistent because of its strong dependence on hardly controllable parameters, such as surface imperfections, nanostructure geometry variations, crystalline structure and others.

Kerr-type nonlinearities, which are predominately determined by heating of the electron gas (its redistribution in a conduction band), strongly depend on the energy deposited by the light in the electron plasma and, therefore, are strongly dependent on the pulse energy and duration in terms of time-response and their value, covering several orders of magnitude: $|\text{Re}(\chi^{(3)})|$ spans the interval from 10^{-18} to $10^{-14} \text{ m}^2/\text{V}^2$ and $|\text{Im}(\chi^{(3)})|$ from 10^{-20} to $10^{-15} \text{ m}^2/\text{V}^2$ for different pulse durations and excitation wavelengths [48]. It also may depend on the size of the particles and thickness of a metal layer due to modification of the electron scattering rate in nanostructures [49–51].

Both coherent and Kerr-type nonlinearities of metals may originate from several families of physical effects. One is related to the nonlinear dynamics of the free carriers' plasma [52–54]. Another important contribution comes from heating of the electron gas and the interband electronic excitation, when the driving field excited electrons from the valence to conduction band of a metal [55]. The latter produces a strong nonlinearity due to a high efficiency of the absorption process but has a spectral range limited to the interband optical transitions. Recent experimental reports and theoretical approaches, relevant to those nonlinear mechanisms will be reviewed hereafter.

8.3 Nonlinearities in Plasmonic Nanostructures

8.3.1 Free-Electron Coherent Nonlinearities in Metals: Hydrodynamic Model

8.3.1.1 Nonlinear Dynamics of Free Electron Gas in Hydrodynamic Description

Various tools for analysis of nonlinear dynamics in plasmas are well established [56], and similar approaches can be adopted for treatment of nonlinear effects in plasmonic structures. In particular, nonlinearity of noble metals can be described using a hydrodynamic-type model describing the electron plasma by means of a charged fluid. This approach was shown to give a qualitative description of the nonlinear interaction of the electron gas with light [57–59], automatically taking into account the electron response defining the metals linear dielectric permittivity. Apart from electric field and damping forces, defining the linear Drude behavior, the hydrodynamic equations include phenomena lacking in the Drude description which treats electrons as standalone charged particles in a periodic potential and encapsulating all solid state aspect of interaction within one effective damping coefficient [60]. Particularly, the hydrodynamic model includes convective acceleration, Lorenz force, and, in the advanced description, a quantum pressure contribution, giving rise to nonlocal effects. Each of these terms leads to a nonlinear behavior due to a varying electron concentration and velocity under the stimulus of an electromagnetic wave. This ultimately defines the nonlinearity of optical properties of metals under visible and infrared light illumination, away from the spectral range of interband transitions. In particular, predictions of second-order responses [52, 61] and Kerr-type nonlinearities [62], based on the hydrodynamic method, qualitatively agree with existent experiments, obtained for pump wavelengths away from interband transitions. In fact, the hydrodynamic treatment was first used by Ritchie in his seminal work [63], where he predicted the existence of surface plasmons. Higher-order corrections and introduction of additional terms, such as quantum pressure of the electron gas and its viscosity, may lead to spatial dispersion contributions and a temperature dependence of the appropriate optical constants [61]. The introduction of conservative ponderomotive potentials may lead to third-order plasma-like nonlinearities of the electron gas [62]. Hydrodynamic model combined with additional Lorentzian resonance terms designed to describe the contributions from the interband transitions is capable to reproduce metal susceptibilities over the entire spectral range [59]. The mesoscopic hydrodynamic approach is integratable with electromagnetic modelling, enabling studies of large scale electromagnetic systems with nontrivial geometries. It is worth noting that ab initio microscopic models are hardly extendable beyond bulk materials and flat surfaces descriptions due to the enormous computation complexity involved [64].

The interaction of electromagnetic fields with objects made from arbitrary nonmagnetic materials is described in terms of the induced polarization $\mathbf{P}(\mathbf{r}, t)$ via the wave equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}, t) + \frac{1}{c^2} \partial_{tt} \mathbf{E}(\mathbf{r}, t) + \mu_0 \partial_{tt} \mathbf{P}(\mathbf{r}, t) = 0, \quad (8.3)$$

where $\mathbf{E}(\mathbf{r}, t)$ is the electric field, c is the speed of light in vacuum, and μ_0 is the vacuum permeability. In general, the spatio-temporal polarizability holds all the information on both linear and nonlinear responses of the material, also including its chromatic dispersion. The polarizability of metal structures can be introduced in this equation via natural hydrodynamic variables: the macroscopic position-dependent electron density $n(\mathbf{r}, t)$ and velocity $\mathbf{v}(\mathbf{r}, t)$, which are subsequently related to the polarization as

$$\partial_t \mathbf{P} = \mathbf{J} = -en\mathbf{v}. \quad (8.4)$$

On the other hand, the dynamics of the free electron gas is determined by a set of hydrodynamic equations [61]

$$m_e n (\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v}) + \gamma m_e n \mathbf{v} = -en(\mathbf{E} + \mathbf{v} \times \mathbf{H}) - \nabla p, \quad (8.5)$$

$$\partial_t n + \nabla \cdot (n \mathbf{v}) = 0, \quad (8.6)$$

where m_e and e are the electron mass and charge, respectively, γ is the effective scattering rate, representing optical losses in a phenomenological way, and

$$p = (3\pi^3)^{2/3} \frac{\hbar^2}{5m_e} n^{5/3} \quad (8.7)$$

is the quantum pressure term evaluated within the Thomas-Fermi theory of an ideal fermionic gas. Equations (8.3–8.6) are inherently nonlinear and provide a self-consistent formulation of nonlinear optical processes originating from free conduction electrons in plasmonic systems. In the perturbative regime of interaction (weak pump field), the leading nonlinear polarizability is of the second order and is the result of the convective acceleration term $\mathbf{v} \cdot \nabla \mathbf{v}$, the magnetic component of the Lorentz force $-e\mathbf{v} \times \mathbf{H}$, quantum pressure and the $n\mathbf{v}$ terms [52]. The further increase of the peak power of the pump pulse brings higher-order nonlinear terms into consideration, resulting in the intermixing of bulk and surface nonlinear effects. Recently, attosecond pulse generation from metallic structures was predicted by applying ultra-strong electromagnetic fields, changing the quantum wavefunctions of electrons [65].

8.3.1.2 Second-Order Nonlinearity of Metals

In the search of the dominating nonlinear processes described by the hydrodynamic model, a perturbative approach can be applied, assuming small nonlinear corrections to the hydrodynamic variables. Under such illumination the hydrodynamic variables of the free electron gas in the nanoparticle can be written as

$$n(\mathbf{r}, t) = n_0(\mathbf{r}) + n_1(\mathbf{r})e^{-i\omega t} + n_2(\mathbf{r})e^{-2i\omega t} + \dots, \quad (8.8)$$

$$\mathbf{v}(\mathbf{r}, t) = \mathbf{v}_1(\mathbf{r})e^{-i\omega t} + \mathbf{v}_2(\mathbf{r})e^{-2i\omega t} + \dots, \quad (8.9)$$

reflecting the fact that due to the inherent nonlinear nature of the hydrodynamic equations, higher harmonics will appear. Through the polarisation term $\mathbf{P}(\mathbf{r}, t)$ taking into account the current $\mathbf{J}(\mathbf{r}, t) = \partial\mathbf{P}(\mathbf{r}, t)/\partial t$ and $\mathbf{J}(\mathbf{r}, t) = -en(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t)$, the terms $n(\mathbf{r}, t)$ and $\mathbf{v}(\mathbf{r}, t)$ will generate higher harmonics in electric and magnetic fields

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_1(\mathbf{r})e^{-i\omega t} + \mathbf{E}_2(\mathbf{r})e^{-2i\omega t} + \dots, \quad (8.10)$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}_1(\mathbf{r})e^{-i\omega t} + \mathbf{H}_2(\mathbf{r})e^{-2i\omega t} + \dots, \quad (8.11)$$

Substituting these expressions in (8.4–8.6), ignoring quantum pressure term in the first approximation and keeping the leading perturbation order, the Maxwell's equation $-\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_b}(n - n_0)$ only allows the second-order nonlinear polarization

$$\mathbf{P}_{NL}^{(2\omega)} = i \frac{e\epsilon_b\omega_p^2}{4m_e\omega^3} \mathbf{E}_1 \times \mathbf{H}_1 + \frac{e\epsilon_b\omega_p^2}{4m_e\omega^4} (\mathbf{E}_1 \cdot \nabla) \mathbf{E}_1 + \frac{ee_b}{2m_e\omega^2} (\nabla \cdot \mathbf{E}_1) \mathbf{E}_1, \quad (8.12)$$

where $\omega_p^2 = n_0 e^2 / (m_e \epsilon_b)$ is the electron plasma frequency, ω is the angular frequency of the fundamental field, n_0 is the concentration of the unperturbed electron gas, m_e is the electron effective mass, and ϵ_b is the background permittivity. It should be noted that the term $\mathbf{E}_1 \times \mathbf{H}_1$ in (8.12) and the nonlinear polarisation as a whole can be expressed only in terms of the electric field and its derivatives [52].

8.3.1.3 Kerr Nonlinearity of Metals from the Hydrodynamic Model

Kerr nonlinearity coefficient can also be directly derived from the hydrodynamic equations. The physical interpretation of the effect is adopted from plasma physics and relies on the introduction of conservative ponderomotive potential. Introduction of the effective potential enables reformulation of the problem of collective interaction of the electron gas and adoption of a single electron description [66]. The nonlocal ponderomotive (Gaponov–Miller) force [67] can be derived from the electron motion in non-uniform electromagnetic field and is given by

$$\mathbf{F}_{\text{PM}}(\mathbf{r}) = -\frac{1}{m_e} \left(\frac{e}{\omega}\right)^2 (\mathbf{E}(\mathbf{r}) \times (\nabla \times \mathbf{E}(\mathbf{r}))) + \mathbf{E}(\mathbf{r}) \cdot \nabla \mathbf{E}(\mathbf{r}). \quad (8.13)$$

The corresponding ponderomotive potential is then defined as

$$\Phi_{\text{PM}}(\mathbf{r}) = e^2 |\mathbf{E}(\mathbf{r})|^2 / 2m_e \omega^2. \quad (8.14)$$

As can be seen both from the above expression and from detailed analysis of electron motion in a single particle representation, electrons will be pushed away from the region of the high field intensity. Consequently, electron concentration at the regions of high intensities will be reduced and, as the result, the local permittivity, which can be obtained in the framework of the Drude model, will become less negative. The process of the electron plasma dilution is balanced by restoring forces applied on an electron by neighboring carriers. The intensity-dependent permittivity is given then by [62]

$$\epsilon_{\text{PM}}(|\mathbf{E}(\mathbf{r})|^2) = \epsilon_m + \frac{3}{2} \left(\frac{\omega_p}{3\pi^2 \epsilon_0 \hbar m_e} \right)^{2/3} \left(\frac{e}{\omega} \right)^4 |\mathbf{E}(\mathbf{r})|^2 = \epsilon_m + \chi_{\text{PM}} |\mathbf{E}(\mathbf{r})|^2, \quad (8.15)$$

where ϵ_m is the linear part of permittivity, and χ_{PM} is the nonlinear ‘‘ponderomotive’’ susceptibility. This Kerr-like coefficient is highly dispersive and for telecom wavelengths is on the order of $10^{-18} \text{ m}^2/\text{V}^2$, comparable with that of nonlinear glasses. Apart from Kerr solitons, this type of nonlinearity could lead to intensity dependent cutoff of SPP modes [62].

8.3.2 *Kerr-Type Nonlinearity of Metals: Two-Temperature Model*

Light absorbed by metal either due to interband or intraband absorption, e.g., via excitation of surface plasmons followed by their fast (10s femtoseconds scale) decay into hot electron excitations, leads in the first instance to modification of the electron distribution in the conduction band. The electron-electron scattering processes (100s femtoseconds scale) that follow thermalize this nonequilibrium distribution establishing an excited state with the electron distribution at an elevated electron temperature. Electron-phonon scattering (picosecond timescale) converts this electron temperature in the increased material (lattice) temperature. The increase in the electron temperature can easily reach several thousands of Kelvins, but for the ultrashort pulses on femtosecond and picosecond scales with reasonable energies, the material temperature changes are often negligible. During several hundreds of femtoseconds before the electron temperature increase is dissipated through electron-phonon interaction, the optical properties of metal are dramatically

changed, with both real and imaginary parts of the permittivity being modified as they both are determined by the electron temperature.

A metal permittivity can be represented within the random phase approximation that includes the dependence of the electron scattering on both electron (T_e) and lattice (T_L) temperatures [45, 68]. Within this approximation, the intensity-dependent permittivity can be described as the sum of the intraband and interband contributions $\epsilon_M = \epsilon_{\text{inter}} + \epsilon_{\text{intra}}$, which depends on the electron's transitions within a conduction sp -band and from a d -band to the sp -band, respectively. The former term can be reduced to a hydrodynamic Drude-like model as

$$\epsilon_{\text{intra}} = \epsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega\gamma_{\text{intra}}(\omega, T_e, T_L)}, \quad (8.16)$$

where $\omega_p = 6.811 \times 10^{15} \text{ rad s}^{-1}$, $\epsilon_\infty \sim 1$ is the high-frequency limit of the permittivity, and $\gamma_{\text{intra}}(\omega, T_e, T_L)$ is determined by both electron-electron and electron-phonon scattering. For materials for which interband transitions are important, e.g., Au, the interband contribution can be described considering a Lorentzian-type contribution to the permittivity as [69, 70]

$$\epsilon_{\text{inter}} = \int_0^\infty g(\hbar x) \times (1 - f(x, T_e)) \times \frac{K}{x \left(x^2 - (\omega + i\gamma_{\text{inter}}(\omega, T_e, T_L))^2 \right)} dx. \quad (8.17)$$

Here, $g(\hbar x)$ is the joint density of states, which can be approximated assuming a dispersionless d -band and a parabolic conduction band with the energy gap between them E_g with $g(\hbar x) \approx \sqrt{\hbar x - E_g}$ (without this approximation on the shape of the bands, the exact joint density of states should be used [71]), $f(x, T_e)$ is the Fermi-Dirac distribution of thermalized electrons near the Fermi Level at a temperature T_e , K is the wavevector-independent constant which depends on the strength of the interband transition dipole moment, and $\gamma_{\text{inter}}(\omega, T_e, T_L)$ is the quasiparticle (electron-hole) scattering rate, which also depends on the electron and lattice temperatures.

The effective nonlinearity described by this model depends on the excitation and probe pulse durations if they are comparable to the characteristic times of the electron relaxation processes [49]. The interplay between the pulse duration and timescale of different processes determines for how long the pulse interact with the excited electron gas and the values of nonlinearity varies widely: $|\text{Re}(\chi^{(3)})|$ from 10^{-18} to $10^{-14} \text{ m}^2/\text{V}^2$, $|\text{Im}(\chi^{(3)})|$ from 10^{-20} to $10^{-15} \text{ m}^2/\text{V}^2$, strongly decreasing for short excitation pulses [48]. The relaxation time of the nonlinear changes depends on the energy supplied by light to the metal (the scattering rates and electron and lattice heat capacities are temperature dependent), as well as on the geometric and modal structure of the plasmonic system and thus can be controlled by nanostructuring, although in a limited range [72].

8.4 Harmonic Generation in Plasmonic Nanostructures

The investigation of second harmonic generation from nanostructures made from centrosymmetric metals [73] started with studies of SHG from flat or curved surfaces [52, 61, 74, 75] and nanoparticles [76, 77]. The derived theory of the process related to the anharmonic response of the free electron gas based on the hydrodynamic model revealed both surface and bulk contributions and following it substantial amount of work was devoted to determining the relative contribution of these terms for various nanostructures and finding the leading nonlinear mechanisms [78–84]. For small spherical particles, SHG was found to be a superposition of dipolar (originating from the retardation effects) and quadrupolar contributions [85–87], while the importance of higher order multipoles for larger particles was also predicted [84]. Following the experimental observations, major mechanisms for the ordinary dipole SHG contribution, forbidden for an ideally symmetric structure were proposed [88, 89]. Particularly, non-centrosymmetric shapes of the particles were invoked to explain the experimental results [90]. Near-field SHG has been studied with near-field detection and excitation, and has been correlated with surface topography [91, 92].

The strategy was identified of getting a qualitative increase of SHG from centrosymmetric metals by designing nanostructures of asymmetric geometrical shapes [90, 93]. At the same time, it was found that in order to make it efficient, the SHG engineering should rely on a certain selection rule, expressed in a semi-empirical criterion as [94–97]

$$E(2\omega) \propto \iint \chi_{nnn} E_n^2(\omega) E_n(2\omega) dS, \quad (8.18)$$

where $E(2\omega)$ is the SHG field, χ_{nnn} is the surface nonlinear susceptibility, $E_n(\omega)$ and $E_n(2\omega)$ are the fields of fundamental and SH modes, respectively, and the integration is performed over the nanostructure surface. Particularly, this approach was used to optimize SHG from multiresonant coupled nanoantennas [96]. The SHG efficiency can be further increased by obtaining beneficial symmetry of local fundamental field [98]. This “selection” rule was also extended to SHG studies from non-centrosymmetric metasurfaces. Other scenarios, such as SHG from metamaterials with hyperbolic dispersion [99] and backward phase-matching [100] were also considered. Finally, it is important to mention that SHG can be strongly affected by the roughness. In the case of metallic films, this was studied both experimentally and theoretically [101]. The main features in the spatial distribution of SHG were discussed in terms of the roughness-assisted SPP generation as well as the LSP hot spots [102, 103].

Third harmonic generation from metallic nanostructures was initially investigated on the examples of flat surfaces and spherical nanoparticles [104, 105]. The advantage of the plasmonic approach was used to demonstrate THG enhanced by the multipolar resonances either in individual metallic nanostructures [106, 107] or the resonant modes of 2D patterned metal surfaces [108–110]. THG in

metamaterials [59] and other enhancement mechanisms, such as those involving interband transitions [111] and phase-matching in a plasmonic waveguide [112] were also considered.

8.4.1 Nonlinear Coupling of Plasmonic Resonances

The implementation of the perturbative hydrodynamic description allows the derivation of a fundamental criterion for efficient generation of localised second harmonic modes at the nanoscale [27]. In the most general case, an arbitrarily shaped metallic nanostructure is considered with a plasmonic resonance at a fundamental frequency, acting as an excitation field for SHG. The nonlinear polarization charge density, acting as the source for the second harmonic field can be obtained from (8.12) under the quasistatic approximation as

$$\rho_{\text{NL}}^{(2\omega)} = -\frac{e\epsilon_b}{2m_e\omega^2} \left(\frac{\omega_p^2}{2\omega^2} + 1 \right) (\nabla \cdot \mathbf{E}_1)^2. \quad (8.19)$$

Note that the divergence of the electric field is a nonvanishing term only at the particle's boundaries, where it is proportional to the surface charge density. These nonlinear charge oscillations serve as the source of the second-harmonic field.

Using a wave equation for the second harmonic fields (analogous to (8.3)) and representing them as a sum of multipoles provides an expression for the SHG field:

$$\mathbf{E}_2 = \sum_k A_k^{(2\omega)}(t) e^{-2i\omega t} e^{-\gamma_k^{(2\omega)} t} \mathbf{F}_k^{(2\omega)}(\mathbf{r}), \quad (8.20)$$

where $\mathbf{F}_k^{(2\omega)}(\mathbf{r})$ are the normalized spatial mode distributions of the multipoles and $A_k^{(2\omega)}$ are their slowly varying (in comparison to ω) amplitudes. Relating the derived nonlinear charge density to the source SH polarisation ($\rho_{\text{NL}}^{(2\omega)} = -\nabla \cdot \mathbf{P}_{\text{NL}}^{(2\omega)}$), one can obtain a relation between the amplitudes of the SH and fundamental resonances [27]:

$$\sum_k \frac{\partial A_k^{(2\omega)}(t)}{\partial t} \nabla \cdot \mathbf{F}_k^{(2\omega)} \approx i \frac{e}{m_e \omega} \left(A_1^{(\omega)} \right)^2 \left(\nabla \cdot \mathbf{F}_1^{(\omega)} \right)^2 e^{(\gamma_k^{(2\omega)} - 2\gamma_1^{(\omega)})t}, \quad (8.21)$$

where $\mathbf{F}_1^{(\omega)}(\mathbf{r})$ is the normalised spatial distribution of the field of the fundamental mode and $A_1^{(\omega)}$ is its amplitude. Taking into account that $\nabla \cdot \mathbf{F}$ is proportional to the surface charge density σ , projecting (8.21) on a surface dipole density of a k -th resonance τ_k and using the orthogonality of the plasmonic modes $\oint \tau_k(Q) \sigma_m(Q) dS_Q = \delta_{k,m}$ (here the integration is performed over the particle surface,

and $\sigma_m(Q)$ and $\tau_k(Q)$ are the relevant quantities at a point Q [113]), the SHG amplitude can be obtained as

$$\frac{\partial A_k^{(2\omega)}(t)}{\partial t} = -i \frac{e}{m_e \omega e_0} \left\{ \frac{\oint \tau_k^{(2\omega)}(Q) [\sigma_1^{(\omega)}(Q)]^2 dS_Q}{\oint \tau_k^{(2\omega)}(Q) \sigma_k^{(2\omega)}(Q) dS_Q} \right\} e^{(\gamma_k^{(2\omega)} - 2\gamma_1^{(\omega)})t} \frac{S}{V} (A_1^{(\omega)})^2, \quad (8.22)$$

where S is the surface area of a particle and V is its volume. The highest efficiency of the excitation will occur for the modes which maximize the overlap integral in the curly brackets of (8.22). Moreover, the multiplication factor of S/V shows explicitly the proportionality of the nonlinear process efficiency to the surface-area-to-volume ratio of the nanoparticle. Note that the nonlinear interaction here has a purely surface origin, supporting the experimental results reported in [88] and elsewhere.

The conversion efficiency of the nonlinear optical processes is usually linked to the local field enhancement since nonlinear polarizabilities are proportional to a certain power of the driving field. Here, a similar link can be made: high surface charge/dipole density leads to high local electric fields. The “matching” integral of (8.22) may reach high values if these surface functions are spatially overlapped, meaning that the corresponding local fields of first and second harmonics have a significant overlap also. In particular, noncentrosymmetric particles can generate the second harmonic more efficiently. Proper matching of parameters can maximize the spatial overlap integral in (8.22), resulting in more efficient SHG.

To achieve efficient SHG into certain localized plasmon modes, a plasmonic particle should have resonances at both fundamental and second-harmonic frequencies (8.22). We illustrate this in the most straightforward case of a single isolated spheroidal particle in free space, investigating the interplay between its different plasmonic resonances. The boundary integral approach [83] can be used to determine plasmonic resonances as

$$\sigma(Q) = \frac{1}{2\pi} \frac{\epsilon_m(\omega) - 1}{\epsilon_m(\omega) + 1} \oint \sigma(M) \frac{\mathbf{r}_{MQ} \mathbf{n}_Q}{|\mathbf{r}_{MQ}|^3} dS_M, \quad (8.23)$$

where $\epsilon_m(\omega)$ is a frequency-dependent metal permittivity, \mathbf{r}_{MQ} is the vector connecting two points on the particle boundaries: any point (M) with a point of observation (Q), and \mathbf{n}_Q is the normal to the boundary at the point Q , while the integration is performed over the particle boundary. This equation has a purely geometrical formulation of determining the set of the eigenvalues α_n obeying $\sigma(Q) = \alpha_n \oint \sigma(M) \mathbf{r}_{MQ} \mathbf{n}_Q / |\mathbf{r}_{MQ}|^3 dS_M$. Then, the actual resonant frequencies can be calculated using the corresponding equation $(\epsilon_m(\omega) - 1) / (\epsilon_m(\omega) + 1) = \alpha_n$, defined by the actual material dispersion $\epsilon_m(\omega)$. The spatial distribution of surface charges and dipole densities can then be found.

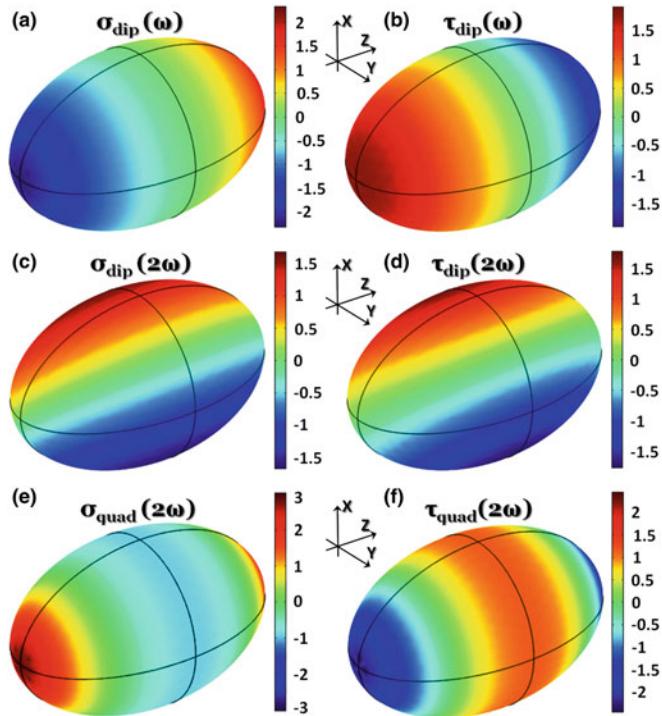


Fig. 8.2 Distribution of surface charge density (left column) and the related surface dipole density (right column) for different plasmonic resonances of the spheroidal particle with the aspect ratio 1:1:1.6 for (a, b) z -polarized dipolar resonance, (c, d) x -polarized dipolar resonance, and (e, f) z -polarized quadrupolar resonance. Reproduced from [27]. Copyright 2012, the American Physical Society

Excitation of a dipolar localized surface plasmon at the fundamental frequency is considered to be the most efficient for the far-field illumination. The surface charge density of the LSP acts as a SHG source and is depicted for a spheroid of an arbitrary aspect ratio in Fig. 8.2a. Figure 8.2d shows the surface density of the second-harmonic dipoles directed along another axis of the spheroid perpendicular to the fundamental dipolar mode, while 8(f) shows this quantity for a co-directed quadrupolar mode. Equation (8.21) predicts a vanishing excitation efficiency for the orthogonal second-harmonic dipole mode, since the surface integral of $\tau_k^{(2\omega)}(Q) \left[\sigma_1^{(\omega)}(Q) \right]^2$ is zero. At the same time, for the quadrupole mode this overlap is nonzero, meaning that it may indeed be excited if its resonance frequency is twice that of the fundamental dipolar mode [27].

The problem now is to match the localized surface plasmon modes at the fundamental and second harmonic frequencies. Two independent geometrical degrees

of freedom (semi-axes of the spheroid) provide sufficient tuning abilities for this case study. Recently, it was shown that higher-order modes (quadrupoles [89] and octupoles [114]) can significantly contribute to the emission pattern of second-harmonic light. Complete analytical solutions for oblate/prolate spheroids exist for dipolar and higher-order modes [115]. Moreover, spheroidal particles have a full analytical solution of the scattering problem, giving scattered field patterns and polarization-dependent cross-sections [116]. Varying the particle geometry, it was found that the matching condition $\omega_{\text{quad}} = 2\omega_{\text{dip}}$ can be achieved [82]. Finally, we would like to note that similar perturbative approach based on surface second-order nonlinearity of metal was used to study SHG-assisted tomography [117] and SHG from hyperbolic plasmonic nanorod metamaterials [26].

8.4.2 SHG and THG from Metallic Nanostructures in Non-perturbative Hydrodynamic Description

Although the perturbative models describe the dominating nonlinear processes in a very illustrative and physically clear way, such analytic treatments are essentially restricted to studies of a predefined nonlinear effect relying on set of analytically-defined electromagnetic modes in a limited number of geometries allowing analytical solutions and at limited (relatively low) excitation powers for which the nonlinear variations is a small perturbation. These conditions significantly limit the capabilities to describe, frequently inter-related, nonlinear processes from arbitrary geometries in strong fields, which can be achieved in plasmonics.

A non-perturbative numerical model for the investigation of nonlinear interactions of light with plasmonic nanostructures can be developed with the use of a time-domain analysis to address the nonlinear dynamics of free electrons without any additional assumptions on the nature of the interaction, which provides the opportunity to explore the interplay between various nonlinear optical processes and geometry of nanostructures [57, 58]. Particularly, this approach gives an opportunity to explore simultaneously both bulk and surface contributions to nonlinear generation processes, as well as the efficiency of sideband generation, involving an interplay between nonlinear effects. A time-domain implementation allows taking all these effects into account by coupling nonlinear hydrodynamic equations, describing the behaviour of the electron plasma, with the Maxwell's equations to model the response to electromagnetic fields. This enables obtaining a universal, self-consistent numerical solution, free from any approximations, allowing for the investigation of nonlinear optical interactions with arbitrary spatially and temporally shaped optical pulses and opening unique opportunities to approach the description of realistic experimental scenarios. Furthermore, the developed formalism paves the way for investigating ultrafast dynamics in mesoscopic and nanoscopic systems with properties defined via microscopic degrees of freedom, which can be introduced in the permittivity model.

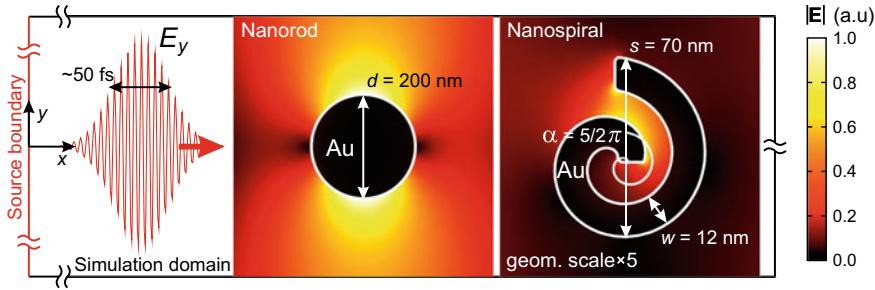


Fig. 8.3 Layout of the time-domain numerical simulations: a ~ 50 fs vertically polarized optical pulse with a central wavelength of 1500 nm is generated at the domain source boundary and illuminates a nanorod or an Archimedean nanospiral from the left. Geometrical parameters of the nanorod and nanospiral are also shown. Adapted from [58]. Copyright 2016, Macmillan Publishers Limited

It should be noted that a frequency-domain analysis of the hydrodynamic equations also allows deriving surface polarizabilities [85] for a particle under the nondepleted pump assumption, as was discussed in the previous section. The effective nonlinear surface polarizability can also be introduced phenomenologically and related to the experimental data. This approach was intensively used to study the SHG from nanoparticles in the quasistatic limit [86, 118] as well as relying on the extended Mie theory to account for nonlocal and retardation effects responsible for the radiation pattern formation [84, 87]. Within the Rayleigh-Gans-Debye formulation, second-harmonic scattering from cylinders of arbitrary size, orientation, and crystallographic structure was investigated [119]. Alternatively, frequency-domain simulations can be used for arbitrary particle geometries, but restricted to an a priori chosen model for the nonlinear response [63]. In contrast, the time-domain approach can be developed free from any initial approximations.

We compare the existing models to the non-perturbative hydrodynamic approach in the canonical case of metallic nanorods (Figs. 8.3 and 8.4). The nonlinear scattering spectra obtained with the full model clearly shows both SH and TH contributions. Since the THG intensity grows with the third power of the fundamental intensity and the SHG does so with the second power of the fundamental intensity, the former shows a comparably faster growth with the excitation intensity (Fig. 8.4a). In the case of the nanorods much smaller than the wavelength, the SHG intensity scales with size as $\propto d^4$ [119], and provides 5 orders of magnitude difference in the SHG intensities from the nanorods with diameters $d = 200$ nm and $d = 12$ nm. Such drastic difference is the result of much smaller dipolar and quadrupolar moments excited at the SH frequency for the smaller nanorod.

The SHG emission diagram from a nanorod has two radiation lobes pointing predominantly in the direction of the driving field and showing the appearance of a modified dipole radiation (SHG dipolar plasmonic resonance along the beam direction is excited due to the retardation effects). In the chosen geometry, the main

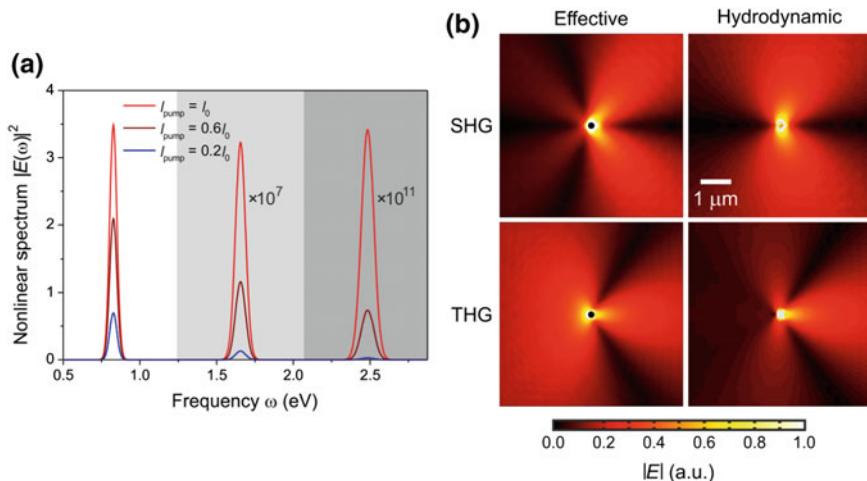


Fig. 8.4 **a** Nonlinear scattering spectra from the infinitely long Au cylinder of a 200 nm diameter simulated for a ~ 50 fs excitation pulse. **b** The distributions of SHG (top row) and THG (bottom row) fields simulated with the phenomenological effective (left column) and microscopic hydrodynamic (right column) models. The colour scale is internal for each plot. Adapted from [57]. Copyright 2015, American Chemical Society

contribution to the SHG can be identified from the convective acceleration and the Lorenz force having comparable contributions, while the quantum pressure effects were not observed. The relative contribution of the different terms may however differ depending on a particular geometry of nanostructures.

To compare this radiation pattern to that obtained from a frequency-domain phenomenological model, Fig. 8.4 shows also the SH radiation pattern evaluated using a two-step model, in which the fundamental field inside the particle is calculated in a first step and the nonlinear field distribution is subsequently derived in a second step using a source of the surface nonlinear polarization $P_{\text{surf}, \perp}^{2\omega} = \chi_{\text{surf}, \perp\perp}^{(2)} (E_{\perp}^{\omega})^2$, where \perp stands for the local normal to the surface [90]. This assumes the nondepleted pump regime, relies on a quasi-Fourier separation of the harmonics, and makes the explicit restricting assumption on the interaction nature to originate from local boundary terms only. The latter follows from the centrosymmetric nature of considered particle's material, so that the dipolar SH radiation from the bulk is forbidden. In this case, one can clearly observe four lobes determined by a quadrupole-like emission added to the dipolar (two lobes) contribution. The phase relations between the two contributions determine the directivity of the total SH emission. One can see the similarities of the phenomenological and non-perturbative models (cf. top fieldmaps in Fig. 8.4); however, the ratio between the dipolar and quadrupolar contribution is heavily distorted in favour of the latter in the phenomenological model. The phenomenologically defined surface polarization term also results in an unphysical singularity at the boundary,

failing to describe the near-field distribution of the nonlinear source; conversely, the time-domain method reproduces the physics of surface interactions with much better accuracy in both near- and far-field regions.

Bulk third-harmonic generation in metals has been considered by the introduction of a nonlocal ponderomotive force, acting on electrons subjected to an electromagnetic field gradient [62]. Theoretical approaches based on the Sommerfeld free-electron models involving solution of the Schrödinger equation in the Kramers–Henneberger accelerating frame describe higher-order harmonic generation at the boundaries [120] and show a good agreement with experimental results [121]. However, the impact of complex geometries on higher-harmonic generation is difficult to evaluate with the latter approach [57].

The TH radiation pattern generated by a nanorod (Fig. 8.4b) possesses strong beaming characteristics in the forward scattering direction as expected for large particles, for which retardation effects lead to constructive interferences in the forward direction only, lowering the backward scattered intensity [122]. In order to perform a comparison between the time-domain approach and the phenomenological model, the TH intensity distribution was simulated with the above-described two-step model using a nondispersive, bulk third-order susceptibility. The radiation pattern derived phenomenologically has similar features to the microscopic model with a stronger backward THG (Fig. 8.4b). The same as for the SHG pattern, strong differences are observed in the near field of the nanocylinder. Since the phenomenological model assumes a third-order susceptibility to be homogeneous across the nanorod, the differences between models may be indicative of a position-sensitive effective third-order susceptibility arising in the hydrodynamic description. While higher harmonic generation is also described by the same hydrodynamic model, sufficient numerical accuracy of the simulations is needed to observe their presence.

8.4.3 Coherent Nonlinear Effects and Nonlocality

Another very important feature of the hydrodynamic description is its inherent ability to describe nonlocal electromagnetic effects. The hydrodynamic nonlocality is a typical example of strong electron–electron interactions between quasi-free electrons of the metal plasma and was proven to describe a variety of phenomena governing the optical response of small plasmonic structures [16]. (Please note that nonlocality discussed here is of different nature than nonlocal effects due to spatial dispersion in a metamaterial [17], which are important for Kerr-type nonlinear response of metamaterials discussed below). In the linear optical regime, the quantum pressure term in the hydrodynamic model is responsible for the appearance of nonlocal effects as it contains a spatial derivative in the linearized model [123]. The effects of electron cloud spillage outside the nanoparticle geometric boundaries can also be taken into account in straightforward way in the

hydrodynamic model by considering electron gas in a finite-height potential well [16, 124]. The relative contribution of nonlocal effects to nonlinear optical properties depends on the characteristic sizes of nanostructures.

The effect of the quantum pressure can be seen from comparison of nonlinear response of plasmonic nanorods of different diameters or Archimedean spiral shaped nanostructures (Figs. 8.3 and 8.5a). Spirals have no symmetry of any kind and, hence, are good candidates for nonlinear optical interactions as they do not obey any geometrical selection rule [125]. Initially, we consider a nonresonant excitation when the excitation frequency is lower than the lowest plasmonic resonances of both nanorods and nanospirals, in order to avoid the influence of resonant effects. The unique ability to either include or exclude the quantum pressure term in the numerical model enables investigating the impact of nonlocality on the nonlinear generation. For large cylinders of 200 nm diameters (blue solid and dashed lines in Fig. 8.5b), the nonlinear scattering intensity (with linear scattering field subtracted) shows a clear signature of higher harmonics up to the 3rd order, though no significant impact from nonlocality. For smaller cylinders of 12 nm in size, the role of the quantum pressure is more significant. At the length scale of few nm ($r = 6$ nm in our case), which is smaller than the mean free pass of electrons and becomes comparable with the radius of nonlocality related to the electron Fermi wavelength (~ 0.5 nm), the nonlocal response starts playing an important role in the nonlinear scattering from the nanostructures. While the structure of the local and nonlocal spectra remains almost unchanged up to the 3rd harmonic (dashed and solid green lines in Fig. 8.5b, respectively), the generated intensity between integer harmonics in the nonlocal case is tremendously enhanced compared with the local counterpart. The effect of nonlocality, however, is much more pronounced in the case of the spiral nanostructure (Fig. 8.5b). For both nanospirals and nanorods, however, the nonlinear signal enhancement rises linearly with the excitation intensity and is related to the nonlocality-induced change in the nanostructures' linear response over the wide spectrum of the excitation pulse. Finally, by adjusting the nanospiral angle α , it is straightforward to match the resonances of the nanospiral to the fundamental frequency (Fig. 8.5c, d). The SHG intensity in this case is dramatically increased by more than 5 orders of magnitude (Fig. 8.5e) [58]. The corresponding effective second-order nonlinear susceptibility of a nanospiral can be estimated to be about $\chi^{(2)} = 600 \text{ pm/V}$ which is 20 times higher than that of lithium niobate.

8.5 Kerr-Type Nonlinearity and Ultrafast Nonlinear Plasmonics

8.5.1 Controlling Light with Light

In the previous sections, metal and dielectric nonlinearities were used to enhance coherent nonlinearities, such as harmonic generation. We now turn our attention to

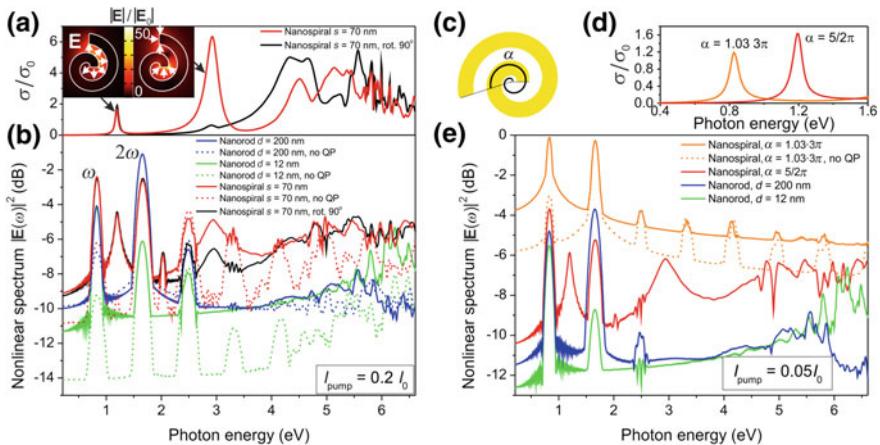


Fig. 8.5 **a** Extinction cross-section spectra of the nanospiral normalized to its geometrical area modelled with linear frequency-domain numerical simulations. The insets show the electric field enhancement maps $|E|/|E_0|$ for the first two nanospiral resonances with respect to the norm of the electric field of the incident wave. The white arrows show the direction of the local electric field. **b** Nonlinear scattering spectra of the nanospirals and nanorods of different parameters for an excitation pulse with a ~ 50 fs duration: solid and dashed lines correspond to the hydrodynamic model with and without the quantum pressure term, respectively. **c** Schematic of the nanospiral indicating the angle α , defining the positions of the nanospiral resonances. **d** Spectra of the extinction cross-sections of the nanospirals with $\alpha = 5/2 \cdot \pi$ and $\alpha = 1.03 \cdot 3\pi$ normalized to their geometrical area ($s = 70$ nm and $w = 12$ nm). **e** Nonlinear scattering spectra of the nanospirals with $\alpha = 1.03 \cdot 3\pi$ (the resonance at the fundamental frequency) and $\alpha = 5/2 \cdot \pi$ and the nanorods with $d = 200$ nm and $d = 12$ nm. Adapted from [58]. Copyright 2016, Macmillan Publishers Limited

Kerr-type third-order nonlinearities [1] and consider how they can be employed in plasmonics to facilitate controlling light with light. The latter ability is required for all-optical signal processing in integrated photonic circuits and its applications in optical communications [126].

In the presence of Kerr-nonlinearity (8.2), the transmission, reflection or absorption of light at a given wavelength (signal light) can be affected by the presence of the strong control light, which induces the nonlinear permittivity changes (the so-called cross-modulation). The signal can be either free-space light or any waveguided mode, including SPP. Alternatively, in self-modulation realization, light can experience self-induced nonlinear propagation effects, if its intensity is strong enough to induce nonlinear Kerr-effects. The latter can manifest themselves as nonlinear refraction, when refractive index of material is changed or nonlinear absorption or transmission if the imaginary part of refractive index is affected.

The Kerr nonlinearity of metals is very fast and, in different regimes, range from tens of fs to few ps, depending on which electron plasma relaxation processes can be accessed with a particular nanostructure [127–129]. Alternatively, nonlinear

optical Kerr-effect can be achieved in dielectrics with slow molecular motion mechanisms (ms timescales), such as liquid crystals and photochromic molecules [130, 131], or fast electro-optic effects as in LiNbO₃, and in semiconductors with fast electronic nonlinearities, up to ps timescales [132]. Through the enhancement of the local fields, plasmonic nanostructures enhance nonlinear effects originating from the Kerr-type nonlinearities either in an adjacent dielectric or a metal itself, and, therefore, promote highly-efficient all-optical modulation, switching and achieving optically tunable photonic properties. These nonlinear effects can be detected, depending on the particular configuration, via observation of either far-field scattering associated with LSPs or modification of guided SPPs [11].

Plasmonic-enhanced light interaction with light can be achieved with metal nanoparticles [133–135] and bulk materials (e.g., nonlinear polymers) doped with metallic nanoparticles [55, 136, 137]. The effective nonlinear susceptibility in this case is determined by the field enhancement at the control light wavelength. The excitation of such composites at the wavelength of the nanoparticle's localized surface plasmon ($\omega_c = \omega_{\text{LSP}}$) leads to an increase in the effective nonlinear susceptibility compared to the off-resonance excitation as described in Sect. 8.2.2. Control illumination induces the local changes of the dielectric permittivity of the nonlinear optical material, which leads to the alteration of the plasmonic resonances of individual plasmonic nanostructures or composites, and, therefore, modifies the propagation of the signal light. Under such conditions, the changes in the real part of the refractive index result mainly in the shift of the plasmonic resonance, while the changes in the imaginary part result mainly in the modification of the resonance magnitude (absorption modulation). This approach can be used in both self-modulation mode, when light propagation is influenced by the changes of its own intensity, or cross-modulation mode, when light at one frequency (control) influences propagation of the light at a different frequency (signal) [11]. Similarly, introducing a nonlinear dielectric material into the arrays of plasmonic nanostructures, e.g., nanosphere or nanorod assemblies, can lower the control light intensity needed for the required level of the signal light modulation [138, 139].

Recently, specifically designed plasmonic nanoantenna resonances have been used to controlling light scattering using the ITO free-carrier nonlinearity in the picoseconds regime [140]. Schemes for utilizing gold nonlinearities and semiconductor-loaded plasmonic nanostructures have also been proposed [141, 142]. The utilization of the field enhancement effects associated with localized surface plasmon modes has also been shown to enable controlled photon tunneling through nanoscale pinholes in a metal film covered with nonlinear polymer [143, 144]. In particular, measurements of photon tunneling through individual, naturally occurring, nanometer scale pinholes have provided indication of “photon blockade” effects, similar to Coulomb blockade phenomena observed in single-electron tunneling experiments [143]. The observations of photon tunneling being gated by light at a different wavelength have also been reported with similar but somewhat larger pinholes [144].

The Kerr-type nonlinear change was used to realise ultrafast all-optical modulation of the transmission through plasmonic nanostructured surface [145], metasurfaces [146], gratings [147], nanoparticle arrays [148], plasmonic cavities [149], and individual nanostructures [135]. Applications of plasmonically-enhanced nonlinearities for pulse polarisation control [150] and shaping [151] were also proposed. The all-optical control of SHG and THG from metallic nanostructures driven by the Kerr-type nonlinearity was experimentally realised showing significant variation of SHG intensity (up to 20%) with control light illumination [47].

8.5.2 *Controlling SPP Modes with Light*

Presently, the implementation of SPPs as signal carriers allows taking advantage of their sensitivity to the properties of nanostructured surface along which they propagate [4] and offers a feasible solution to the problem of all-optical control of light. Particularly, the change in the SPP propagation can be achieved via small changes of the refractive index in a nonlinear optical material deposited on top of the nanostructured plasmonic surface. This effect was first shown in the case of self-induced changes for SPPs excited on smooth metal films in the Kretschmann configuration, utilizing phase transformations in a liquid crystal under the influence of SPP-induced thermal effects [131]. Both switching of the reflected light and bistability (explained by a positive feedback due to the intensity dependent refractive index of the liquid crystal) with the intensity of the incident light was observed [11]. A highly-efficient method of modulation of SPP waves using structural phase transformation in Ga was demonstrated theoretically and realized experimentally in [152] and [153, 154], respectively (Fig. 8.6a, b).

In order to optically modulate and switch SPP signals in a waveguiding geometry two approaches are possible based on the control light induced changes of the real and/or imaginary part of the permittivity of the metal or adjacent dielectric [11]. The latter approach was implemented with light-induced absorption modulation in the dielectric [38, 39] and metallic [152, 155] components of the plasmonic waveguide. Particularly, the modulation of the SPP wave propagation at nanosecond timescale was experimentally demonstrated via light-induced structural (phase) changes in Ga, with two phases having significantly different optical properties [153, 154]. As in either phase Ga has significant losses in comparison with traditional plasmonic materials, such as gold and silver, a switching Ga section from one phase to another in a conventional SPP waveguide offers a feasible solution for the realisation of the switch in the integrated geometry. In the case of intrinsic Au nonlinearity, the excitation of nonequilibrium electrons via intra- and inter-band transitions leads to small changes of $\text{Re}(\epsilon)$, but very significant changes of $\text{Im}(\epsilon)$, allowing ultrafast all-optical light modulation at a picosecond scale [127, 155]. Integrated geometry has also been proposed with a signal SPP absorption modulation and stimulated emission under the influence of a control SPP beam at the different frequency [33].

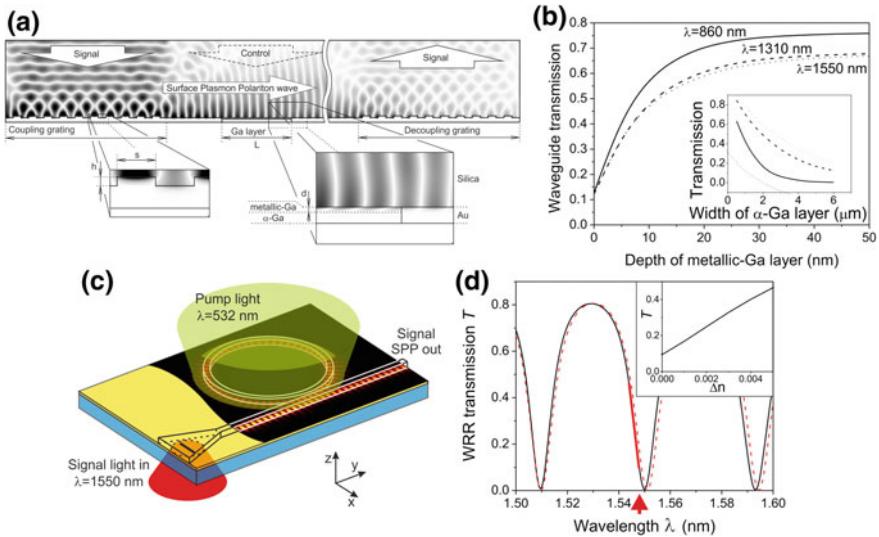


Fig. 8.6 **a** SPP gold-on-silica waveguide containing a gallium switching section. The plasmonic film is at the bottom of the silica substrate. Field mapping shows the magnitude of the transverse component of the magnetic field. **b** Waveguide transmission as a function of the depth d of the metallic gallium layer. Length of the Ga section: 860 nm–2.2 μm, 1310 nm–5.8 μm, and 1550 nm–9 μm. Inset shows the dependence of transmission on the width L of the gallium section. Adapted from [152]. Copyright 2004, The American Institute of Physics. **c** Schematics for modulation of an SPP mode by light using a waveguide ring resonator component. Incorporated field-map is the simulated $|\text{Re}(E_z)|$ cross-section of the SPP mode at the distance of 10 nm above the metal surface. Adapted from [34]. Copyright 2011, The Optical Society of America. **d** WRR transmission as a function of the wavelength: (solid black line) numerical simulations, (dashed red line) analytical fit. Inset shows the dependence of the WRR transmission on the change of a refractive index of the ring material

The variation of the real part of the metal or dielectric permittivity influences the phase velocity of an SPP wave, and thus its phase at the waveguide output [11]. In order to convert this to intensity modulation, various phase-sensitive configurations are used such as Mach-Zehnder interferometers (MZIs) or waveguide-ring resonators (WRRs) [34, 35, 156, 157]. The signal modulation is achieved with a MZI by inducing a phase change of the SPP in one of its arms and interfering it with the wave from the other, unmodulated arm. In the case of a WRR, the induced changes of the refractive index of the ring lead to the modification of its resonance conditions and provide the modulation of the SPP mode in the linear waveguide (Fig. 8.6c, d). Both MZIs and WRRs can be realized on various types of plasmonic waveguides, such as metal-insulator-metal, V-groove or dielectric-loaded waveguides. Figure of merit introduced for the WRR components shows that plasmonic-based WRRs are intrinsically better than similar photonic ones due to the stronger changes in the effective refractive index of the mode induced by the changes of the refractive index of constituting material ($\partial n_{\text{eff}} / \partial n$), resulting from

higher mode localisation and field-enhancement effects [14]. The dielectric refractive index variation of 0.001 provides 50% change in the nonlinear polymer loaded plasmonic WRR transmission with only 5.5 μm ring radius for operation at telecom wavelengths (see inset to Fig. 8.6d). Even higher integration density of the active components can be achieved implementing a high refractive index nonlinear material, such as LiNbO₃. The radius of the WRR ring in this case can be as small as 250 nm with strong mode confinement and acceptable radiation losses [36].

Thermal and photochromic nonlinearities providing very strong changes in the real part of permittivity occur at very long (ms) time scales while for majority of photonic applications much faster nonlinearities are needed, though faster nonlinearities have generally smaller nonlinear susceptibility. In the latter case, the resonant regime of operation of nonlinear components is required where smaller changes of the refractive index can lead to significant changes of the optical response. The requirements on operational speed and energy consumption may additionally favor the use of active plasmonic components in photonic integrated circuits [36].

8.5.3 *Nonlinear Plasmonic Crystals and Optical Bistability*

Nanostructured metal surfaces such as surface polaritonic crystals (SPPCs) [4, 158] provide much more flexibility than the planar surface waveguides for tailoring plasmonic resonant conditions and the electromagnetic field enhancement in devices with subwavelength thickness and a size of just a few wavelengths [11]. SPP Bloch modes supported by SPPCs can be excited without any special arrangements, as the required momentum matching between SPPs and photons is provided by the periodic structure. Introducing nonlinear dielectrics in such nanostructures allows harnessing of the field enhancement effects to observe nonlinear effects and bistable behavior at low light intensities. The dispersion of the SPP Bloch modes of a SPPC can be engineered by varying its period and a unit cell geometry, which leads to the control over both the resonant spectral positions of the plasmonic modes and their propagation characteristics. In particular, the realization of SPP modes with a low group velocity results in high field enhancement, which is beneficial for the nonlinearity enhancement. Moreover, with a flat SPP dispersion, it is much easier to observe the shift of the resonance [159] if the nonlinear effects are induced. Since all optical properties of the SPPCs, reflection, absorption and transmission, are determined to large extent by the their Bloch mode band structure, the all-optical control of the latter via the modification of the refractive index of metal or adjacent dielectric leads to the modulation of the intensity of the signal light reflected from or transmitted through the SPP crystal. Due to the presence of multiple bands in the SPPC dispersion, there is an additional flexibility for simultaneous resonant excitation of SPP Bloch modes at both control and signal

wavelengths and, therefore, additional enhancement of nonlinear interaction between control and signal light.

Integratable all-optical devices can be developed using nonlinear surface plasmon polaritonic crystals. SPP crystals have optical properties similar to two-dimensional photonic crystals but act on surface polaritons rather than “bulk” photons. SPPCs are most often considered with either a periodic arrangement of grooves or stripes (1D SPPC) or as a two-dimensional periodic arrangement of holes or bumps of various shapes (2D SPPC). A fundamental difference between conventional photonic and surface polaritonic crystals comes from a different electromagnetic field distribution close to the surface. Surface polaritons are intrinsically two-dimensional excitations whose electromagnetic field is confined at a metal-dielectric interface. Thus, in contrast to photonic crystals, strong electromagnetic field enhancement takes place at a SPP crystal’s interface related to the surface polariton’s field confinement. This enhancement results in a much stronger effective nonlinear optical response achievable with surface polaritonic crystals as it is related, as described above, to the local field strength $\chi_{\text{eff}}(3\omega) = L^2(\omega_c)\chi^{(3)}(\omega)$, where $L(\omega_c)$ is the field enhancement coefficient at the control light frequency ω_c . This can significantly assist nonlinear optical effects in such structures. The advantage of a surface polaritonic crystal for the development of all-optical active elements is two-fold: a high sensitivity of SPP resonances to minor modifications of optical properties of surroundings, and the electromagnetic field enhancement related to surface plasmon modes that facilitates achieving these modifications via the optical Kerr effect at low illuminating light intensities. As a consequence of the dielectric constant change, the SPP modes will experience a frequency shift, thus modifying the resonant conditions of the SPP mode excitation and related optical transmission or reflection.

Nonlinear optical transmission through plasmonic structures hybridized with nonlinear dielectrics can be observed in a variety of geometries [11]. Two families of the effects can be realized. The first includes self-induced effects, when light induces nonlinear changes in the dielectric and experience self-action as the result of these nonlinear changes. For example the self-action of light was demonstrated for SPPCs hybridized with liquid crystals [160], demonstrating their good intensity-limiting properties. The second family comprises the phenomena when the external control light is used to change the optical properties of the SPPC and, therefore, reflection and transmission of a signal light. For instance, this can be realized through the hybridization of the plasmonic crystals produced by periodic hole arrays in metallic films with nonlinear polymers and utilizing the strong electromagnetic field enhancement of the control light in the cylindrical channels in order to achieve efficient modulation of the light transmission through the SPPC via control of the Bloch modes [161, 162]. In general, nonlinear effects are observed for all transmission resonances related to SPP Bloch waves and cylindrical surface plasmon excitations [161]. At the same time, since the local field enhancement depends on the SPP crystal parameters as well as on the properties of the supported

surface plasmon modes, the range of the realized control light intensities depends on the geometrical and material properties of the nanostructures and, therefore, is a complex function of a plasmonic crystal design. A prominent example of light-induced nonlinear effects is the excitation of transient refractive index grating in a metallic film via free-electron nonlinearity of gold with a consequent control of coupling of signal light into SPPs on sub-ps timescale [155, 163].

Up to 60% change in the transmission facilitated by the SPP Bloch mode was demonstrated in pump-probe experiments. Both increase and reduction of the transmission signal were shown for different resonances (Fig. 8.7), depending if their nature is purely Bloch mode related or is the result of an interplay between the Bloch modes and cylindrical plasmon modes in the holes forming the crystal. The observed SPP band-gap shift corresponds to the average changes of the effective refractive index of the polymer induced by the control light of 10^{-3} – 10^{-4} while the local changes can be higher [158]. The spatial distribution of the local control field on a metal surface and its variations with an intensity of the control light lead to the self-consistent dependence of the spatial variations of the polymer refractive index

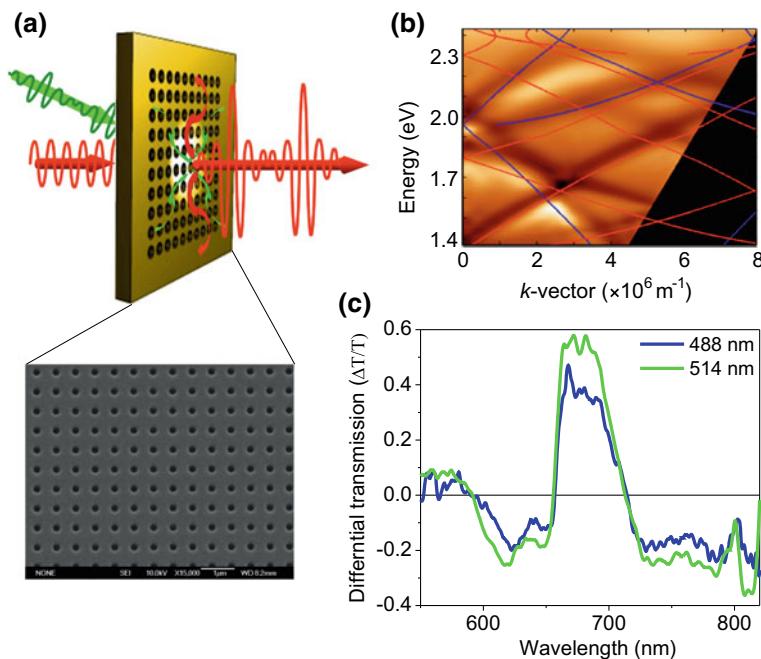


Fig. 8.7 **a** Schematics of modulation of light by light via coupling to surface plasmon polariton modes in a plasmonic crystal. The inset shows an SEM image of the crystal, produced by a square array of circular 200 nm holes with a period of 600 nm milled in a 220 nm thick gold film. **b** SPP band structure of the crystal. **c** The differential transmission induced by 488 nm and 514 nm control light for a crystal in **(b)** covered with a 100 nm thick nonlinear polydiacetylene film. Adapted from [158]. Copyright 2008, WILEY-VCH

on the intensity of the control light and, thus, the changes in the signal transmission. At certain control wavelengths, this allows the observation of the bistable behavior in the signal transmission [162]. Although the modulation rate in the hybrid dielectric-metal structures depends on the relaxation time of both the dielectric material and metal, usually it is mainly determined by the former, as it is generally a slower process.

8.5.4 Nonlinear Plasmonic Metamaterials

Plasmonic metamaterials provide additional opportunities for utilizing nonlinear effects for all-optical switching, since not only the plasmonic resonances of individual nanostructures but the interaction between them in a metamaterial can be influenced [21]. The modification of the refractive index of the embedding dielectric or substrate results in the modification of both individual plasmonic resonances and their interaction, resulting in the enhancement of the nonlinear response. In contrast to plasmonic crystals, metamaterials have characteristic sizes of the elements and separations between them much smaller than the operating wavelength, so that they can be described by introducing effective medium parameters through the averaging over many periods. While such effective medium approach describes the linear optical properties of metamaterials (reflection, transmission and absorption) well, it should be applied to treatment of nonlinear optical properties with caution as it does not take into account the local fields inside the metamaterial composites which can vary significantly [164]. (For the same reason, it also is not applicable for the description of the Purcell effect for the emitters inside the metamaterial [165].) The extension of the effective medium model taking into account local field variation via nonlocal spatial-dispersion effects may alleviate this problem in some cases, in particular in the case of hyperbolic metamaterials [166–168].

The Kerr-type nonlinear effects have been demonstrated with the nanorod-based metamaterials where the spectral position of the metamaterial plasmonic mode was controlled by modifying the refractive index of a polymer layer between the nanorods [138]. Similarly, the dielectric and metal nonlinearity of split-resonator-based metamaterials has been explored. Efficient all-optical control has been demonstrated in the SRR-based metamaterial hybridized with carbon nanotubes used as a nonlinear material [169]. Very fast relaxation time under 100 fs may be achievable for bare gold SRR metamaterials using the intrinsic nonlinearity of gold under two-photon excitation [170]. Efficient all-optical modulation has also been achieved by controlling the coupling strength between molecular excitons and plasmonic excitations in the nanorod metamaterial, which is extremely sensitive to any perturbation of the system, including changes of the metal permittivity [171].

8.5.5 *Epsilon-Near-Zero Metamaterials: Engineering Kerr Nonlinear Response*

Plasmonic metamaterials provide an opportunity to develop a new approach to enhance nonlinearity utilizing the effects arising in the epsilon-near-zero regime [44, 166], when the real part of the effective medium permittivity of the metamaterial is close to zero. In this case, nonlocal spatial dispersion effects mentioned above become important. The nonlocal effects depend strongly on the losses in system and can be significantly modified by controlling loss in the Au nanorod meta-atoms. The modulation based on nonlinear response of Au under the interband excitation leading to the significant changes of $\text{Im}(\epsilon)$ of Au results in a very strong changes of the metamaterial transmission due to modification of the nonlocal response. The transmission changes of up to 80% have been observed with a sub-picosecond response time (Fig. 8.8). To achieve this performance in $100 \times 100 \text{ nm}^2$ devices, 10 fJ pulses are sufficient [44]. Such metamaterials can also be integrated in nanophotonic waveguides, e.g., Si waveguides, to provide very

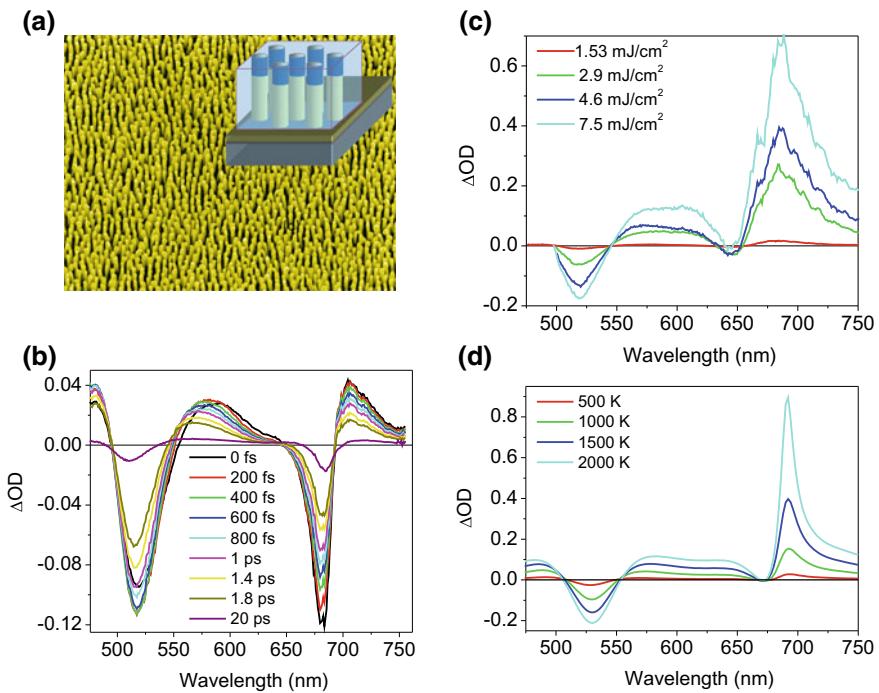


Fig. 8.8 **a** Optical metamaterial based on an array of gold nanorods (diameter 20 nm, length 70 nm) with an average inter-rod separation of 400 nm. **b** Transient extinction spectra for various time delays between the pump and the probe pulses for TM illumination at 20 degrees. Measured **(c)** and simulated **(d)** transient extinction spectra of the nanorod metamaterial for various pump fluencies for TM illumination at 40 degrees. Adapted from [44]. Copyright 2011, Macmillan Publishers Limited

efficient all-optical modulation and switching of the guided signals [46]. In this case, the modulator consists of a hyperbolic metamaterial slab deposited on top of a Si waveguide. The active control over the modulator's transmission is achieved optically by controlling the level of hybridization between the fundamental mode supported by the Si waveguide and extraordinary modes supported by the hyperbolic metamaterial slab. Strong coupling within the modulator corresponds to a mode delocalization in the metamaterial leading to increased insertion losses and a consequent minimum in the modulator's transmission. The operation performance of this modulator is comparable with its integrated all-optical counterparts, but offering much smaller footprint [36].

The strong Kerr nonlinearity of metals is significantly restricted to the spectral range of the interband electronic transitions where efficient electron excitation into a conduction band takes place, leading to the strongest nonlinear response. This nonlinearity becomes weaker at frequencies away from the interband absorption, limiting its usefulness in the (near) ultraviolet spectral range of control light wavelength. On the other hand, the strong absorption near the interband resonance, in many cases, prohibits useful applications when the signal light, which may be controlled by this nonlinearity, is in this spectral range. If, however, Au nanostructures form a hyperbolic-type metamaterial, the nonlinearity near the effective plasma frequency [172] of the metamaterial where the ENZ-regime is achieved (Fig. 8.9a, b) is strongly enhanced [44]. Not only the enhanced nonlinear response but the sign of the nonlinearity at a required wavelength can be engineered by changing the geometrical parameters of the nanorod metamaterial [45].

The z-scan measurements allow direct determination of the nonlinear refraction, γ , and absorption, β , coefficients related to the intensity-dependent, complex effective refractive index of the metamaterial $\tilde{n} = n + i\alpha/(2k_0)$, k_0 being the light wavevector and $n(I) = n_0 + \gamma I$ and $\alpha(I) = \alpha_0 + \beta I$, where n_0 and α_0 are the linear refractive index and absorption, respectively, and I is the intensity of the incident light. As a consequence of the metamaterial's anisotropy, the retrieved effective values of nonlinear refraction $\gamma(\theta)$ and absorption $\beta(\theta)$ coefficients correspond to the variation of the metamaterial's effective linear refractive index for a specific incidence angle θ_i that can be related to the components of the effective third-order nonlinearity tensor of the anisotropic metamaterial as [45]

$$\epsilon_{\text{eff}}(\theta_i) \approx \epsilon_{\text{eff}}^0(\theta_i) + 3\epsilon_s \sin^2 \theta_i |E_0|^2 \left[\frac{\chi_{xxzz}^{(3)}}{\epsilon_{xx}^0} + A(\theta_i) \frac{\chi_{zzzz}^{(3)}}{\epsilon_{zz}^0} \right], \quad (8.24)$$

where $A(\theta_i) = 1 - \frac{\epsilon_{\text{eff}}^0(\theta_i)}{\epsilon_{xx}^0} - \frac{\epsilon_s \sin^2 \theta_i \chi_{xxzz}^{(3)} - \chi_{zzzz}^{(3)}}{\chi_{zzzz}^{(3)}}$ is an angle-dependent but nonresonant term, ϵ_{xx}^0 and ϵ_{zz}^0 are the components of the homogenized linear permittivity tensor of the metamaterial, $\chi_{xxzz}^{(3)}$ and $\chi_{zzzz}^{(3)}$ are the components of the third-order nonlinear susceptibility tensor, dominant for the TM-polarized light, ϵ_s is the permittivity of the medium adjacent to the metamaterial where the incident wave is coming from, and E_0 is the incident electric field amplitude. One can see that near the effective

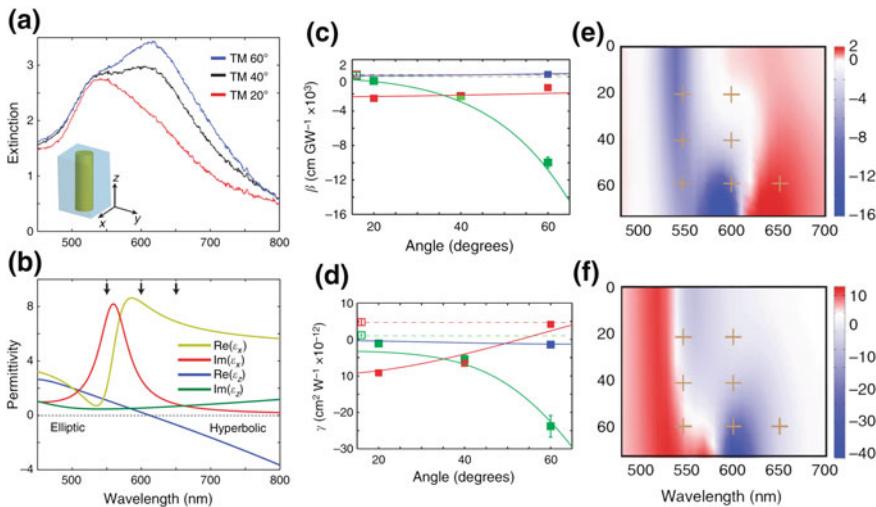


Fig. 8.9 **a** Extinction spectra of the nanorod metamaterial measured at different angles of incidence. **b** Effective permittivities of the metamaterial deduced from the effective medium model. Arrows indicate the wavelengths where the nonlinear coefficients were measured. **c** Nonlinear absorption and **d** nonlinear refractive index dependence on the angle of incidence for different wavelengths: (squares) experiment, (lines) simulations: (red) 550 nm, (green) 600 nm, and (blue) 650 nm. Empty squares and dashed lines are for a smooth Au film. All measurements and simulations are for TM-polarized incident light. The errors bars in **(c)** and **(d)** are smaller than the size of the square unless indicated. Nonlinear absorption **(e)** and refraction **(f)** coefficients for the nanorod array metamaterial calculated using a finite element method (the colour scale units are the same as for y-axis in **(c)** and **(d)**, respectively). The crosses indicate the points where the experimental data are presented in **(c)** and **(d)**. Adapted from [45]. Copyright 2016, Macmillan Publishers Limited

plasma frequency when ϵ_{zz}^0 is close to zero, the effective nonlinearity is greatly enhanced.

The measured and estimated from the two-temperature model nonlinear coefficients are very different for a smooth Au film and the Au-based metamaterial both in the value and sign of the nonlinearity (Fig. 8.9c–f). For wavelengths in the elliptic dispersion regime ($\lambda < 600$ nm), nonlinear absorption of the metamaterial is almost constant, whereas nonlinear refraction which is defocusing for smaller angles, decreases with the incidence angle and changes sign becoming a focusing one. In the ENZ regime with the effective plasma frequency $\text{Re}(\epsilon_z) \sim 0$ at around 600 nm, both β and γ significantly increase with the incident angle. The largest nonlinearity was experimentally measured in these conditions at an angle of incidence of 60° and at the wavelengths close to the effective plasma frequency of the metamaterial: $\gamma \approx -2.4 \times 10^{-11} \text{ cm}^2/\text{W}$ and $\beta \approx -9967 \text{ cm/GW}$. In comparison, the measured nonlinear coefficients for a smooth 50-nm-thick Au film sharply decrease with increasing wavelength away from the interband transitions with the experimentally measured values $\gamma \approx 4.8 \times 10^{-12} \text{ cm}^2/\text{W}$ and $\beta \approx 272 \text{ cm/GW}$ at 550 nm wavelength and $\gamma \approx 1.2 \times 10^{-12} \text{ cm}^2/\text{W}$ and $\beta \approx 122 \text{ cm/GW}$ at 600 nm wavelength,

which are the typical values for Au under femtosecond excitation [48]. At the same wavelengths, γ and β of the nanorod metamaterial are approximately 20 and 100 times larger than those measured for a smooth Au film. Surprisingly, the maximum value obtained for γ and β for the metamaterial away from Au interband transitions is larger than the maximum values measured for a smooth gold film close to the interband transition where they are highest. While in the studied range of frequencies, the nonlinearity of smooth Au is always positive (induced absorption and focusing nonlinearity), the Au nanorod metamaterial can provide either induced absorption or transparency, focusing or defocusing nonlinearity, depending on the combination of a light wavelength and an angle of incidence. Thus, not only the strength but also the sign of the nonlinearity can be designed with plasmonic metamaterials.

The wavelength at which the strongest nonlinearity is observed can be engineered and pre-defined at the fabrication stage by setting the proper geometrical structure with the same constituent materials. Therefore, an artificial optical material with a strong nonlinear response can be realized at the wavelengths where the constituent materials have negligible nonlinearity. As an example, in [45] a nonlinear Au nanorod metamaterial has been designed for the telecommunication spectral range where Au has a negligible nonlinear response.

It should be noted that the condition on the nonlinearity enhancement in the ENZ regime similar to (8.24) is also valid for conventional plasmonic materials. Very strong nonlinearity has been observed for thin ITO films at the plasma frequency in the telecommunication spectral range [173].

8.6 Nonlinear Surface Plasmon Polaritons

Nonlinearity may lead to completely unexpected behavior of SPPs. In a linear case, when considering nonmagnetic media, SPPs are always TM-polarized surface waves with the electric field normal to the conductor interface and parallel to their propagation direction. However, if the Kerr-type nonlinearity is introduced, TE-polarized nonlinear surface waves may exist at the interface between a nonlinear dielectric and a linear metal or a linear dielectric and a nonlinear metal [8]. The former case has more stringent requirements on nonlinearity of a dielectric (it should be defocusing and stronger than nonlinearity of metal), while the latter case can always be achieved considering free-carrier nonlinearity of metals under strong enough intensity of exciting light. Propagation of high-intensity SPPs at the interfaces bounded by a nonlinear medium results also in other interesting phenomena such as soliton-like behavior of plasmonic waves. In this section, several nonlinear phenomena that influence propagation of SPPs will be discussed considering nonlinearities stemming either from the metal or the dielectric components of a plasmonic waveguide.

8.6.1 Nonlinear SPP Modes Due to Ponderomotive Nonlinearity

General nonlinear phenomena in planar SPP waveguides were extensively studied and various combinations of nonlinearities were considered [174]. The propagation of SPP modes is affected by the third order nonlinearity even in the most basic SPP supporting structure: the single metal–dielectric interface [62]. At the wavelengths longer than the interband transitions range, the third-order nonlinearity can be attributed to the previously discussed ponderomotive forces. This dynamic phenomenon, having a direct counterpart at classical plasmas, leads to repelling of electrons from the region of the high field intensity. As the result, the corresponding nonlinear effect is related to the electron depletion in the high-intensity regions that are located just at the metal-air boundary. The linear effective index of SPP mode is given by $\sqrt{(\epsilon_M + \epsilon_d)/\epsilon_M \epsilon_d}$, where ϵ_M and ϵ_d are the permittivities of metal and dielectric media, respectively. As a result of the carriers' depletion next to the guiding boundary, the negative metal permittivity near the interface becomes less negative, and may approach a critical value of $\epsilon_M = -\epsilon_d$. At this condition, the SPP mode approaches its cutoff. For light intensity corresponding to this critical value, the intensity-induced modal reshaping results in equal, but opposite, power flow in the metal and dielectric. The corresponding SPP intensity-dependent nonlinear dispersion relations show the intensity dependent cutoff wavelengths for the nonlinear SPP modes (Fig. 8.10).

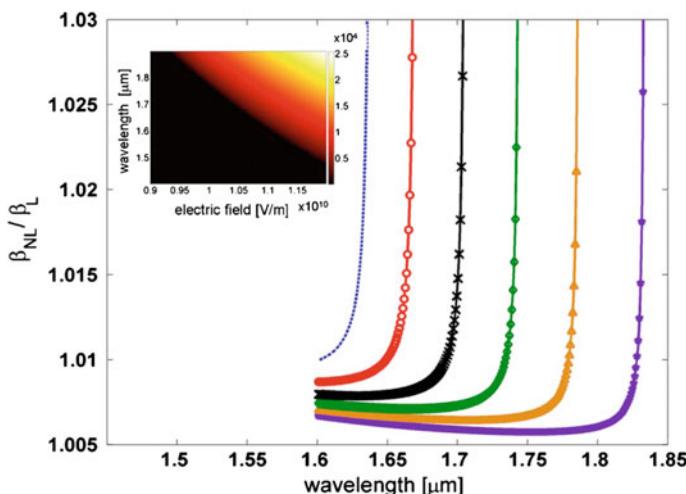


Fig. 8.10 Dispersion of a single-surface nonlinear SPP on air–gold interface for different interface electric field amplitudes: (dashed blue) 12 GV/m, (red circles) 11.5 GV/m, (black crosses) 11 GV/m, (green diamonds) 10.5 GV/m, (brown triangles) 10 GV/m, (purple stars) 9.5 GV/m. The inset shows the nonlinear effective index normalized to the linear one as a function of the wavelength and the field amplitude. Reproduced from [62]. Copyright 2010, The Optical Society of America

8.6.2 Plasmon-Solitons Due to Third-Order Nonlinearity

Diffraction of SPP modes can be suppressed with nonlinear effects leading to their solitonic behavior [175]. Solitons (non-diffracting waves) are one of the most explored features in nonlinear physics and optics, in particular, and can represent both spatial and temporal dispersionless propagation. One of the remarkable features of plasmonic modes is that they can be inherently highly confined beyond the classical diffraction limit [176]. Dielectric-loaded, metal stripe or tapered plasmonic waveguides provide strong confinement of guided modes with even stronger confinement in metal-insulator-metal (MIM) architectures [14, 177, 178]. Generally, if 3D confinement is neglected, nonlinear dispersion of the SPP modes in MIM waveguides with even-order nonlinearities (Kerr or higher) can be analytically derived [179, 180]. When the confinement is taken into account (in, e.g., a Kerr-nonlinear slab bounded by the metal slabs), hybrid vector spatial plasmon-solitons may emerge. The reduction of the gap between the metal layers increases both the linear effective index (n_0) of the fundamental SPP mode and the effective nonlinear coefficient n_2^{eff} [181]. Under these conditions, the lateral confinement of plasmon-soliton modes increases with the reduction of the separation between the metal layers [181], which is the opposite trend compared to a nonlinear all-dielectric waveguide, so that the 3D soliton confinement can be achieved beyond the classical diffraction limit impossible with the dielectric nonlinear waveguides (Fig. 8.11).

Nonlinear propagation of surface plasmon polaritons on the boundary of a metal and a nonlinear Kerr dielectric in presence of losses was further considered and self-focusing phenomenon with the formation of slowly decaying spatial solitons was numerically proved [183]. It is worth noting, that plasmon-solitons are dissipative owing to inherent losses of metal components. In order to overcome the propagation losses, tapered plasmonic waveguides have been proposed [184]. The tapering at properly chosen angle generates additional field confinement enhancing further the field intensity and enabling longer soliton propagation.

The propagation losses may also be compensated by incorporating active medium in a waveguide. Spatial surface-plasmon solitons on the metal stripe surrounded by active and passive dielectric media were considered [185]. Stable plasmon-solitons were achieved by proper coupling between SPPs propagating along the active and passive interfaces. Multilayered metal-dielectric nonlinear composites can also support the so-called discrete solitons. In this case, the discrete diffraction may be suppressed by introducing nonlinearity into one or more material components. Discrete solitons at the nanoscale exhibit considerably different behavior from conventional nonlinear dielectric waveguide arrays due to combined interplay between periodicity, nonlinearity, and SPP-related field confinement [186].

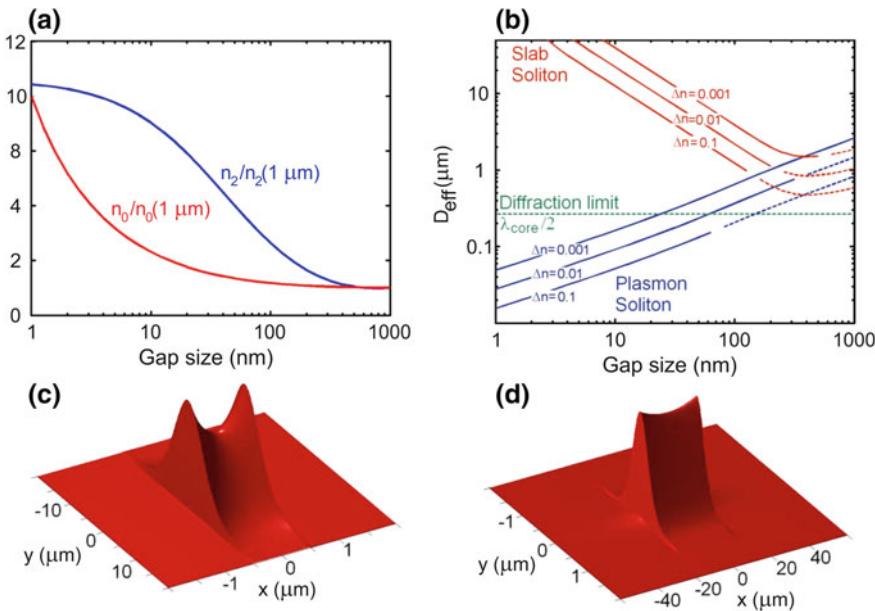


Fig. 8.11 **a** The dependence of the normalized (red) linear and (blue) nonlinear effective indexes of the plasmonic gap mode at $\lambda_0 = 1500$ nm on the thickness of the Kerr-nonlinear slab embedded in silver. **b** The dependence of the effective modal size D_{eff} on the thickness of the Kerr-nonlinear slab embedded in (blue) silver and (red) air for $\lambda_0 = 820$ nm and $n_0 = 1.5$. **c, d** Intensity distributions of the plasmon–soliton beam in the Kerr medium placed in a gap between silver layers: **(c)** $\Delta n = 0.005$, $\lambda_0 = 1550$ nm, $d = 1 \mu\text{m}$, $\epsilon_m = -103.5 - 10i$, $D_{\text{eff}} = 2.4 \mu\text{m}$, and **(d)** $\Delta n = 0.1$, $\lambda_0 = 820$ nm, $d = 30$ nm, $\epsilon_m = -30.2 - 1.6i$, $D_{\text{eff}} = 95$ nm. Reproduced from [182]. Copyright 2007, IEEE

8.6.3 Cascaded Plasmon-Solitons Due to Second-Order Nonlinearity

Second-order nonlinearity can lead to spatial soliton wave formation via the effect of second-harmonic generation, as was theoretically predicted [187] and experimentally demonstrated in, e.g., potassium titanyl phosphate (KTP) crystals [188] and planar LiNbO₃ waveguides [189]. The principle behind such spatial solitons is the collinear propagation of two beams at both fundamental and second-harmonic frequencies. These beams exchange their energies via the second-order polarizability, which coined the term ‘cascaded $\chi^{(2)}$ solitons’. Such an exchange provides the maximum phase delay at the region of high intensity, resulting in self-focusing. This nonlinear phenomenon provides many opportunities for applications and fundamental studies of solitonic effects [190].

Here, following [191], we illustrate the concept of cascaded $\chi^{(2)}$ SPPs propagating at the interface between a linear dielectric and a metal with the nonlinearity described by the hydrodynamic model. Two co-propagating SPP beams are considered at ω and 2ω frequencies, which are nonlinearly coupled through the nonlinear polarisation of the metal.

SPP modes are strongly confined to the metal–dielectric interface where the nonlinear interactions take place. The mismatch between the effective refractive indices of the fundamental and the second-harmonic beams, reflecting the dispersion of the SPP waves and determining the essential phase relations between the waves, should be taken into account. The finite propagation distance of SPPs, which is determined by Ohmic losses in the metal, also plays a significant role in the formation of this type of solitons.

The SPP beams at the fundamental and second-harmonic frequencies propagate collinearly along the z -axis with the transverse field profile described by the Gaussian distributions. Comparing the evolution of the SPP profiles during the propagation in linear and nonlinear regime, the formation of the solitons can be observed (Fig. 8.12). The intensity distributions $|E_{1,2x}|^2$ obtained in the linear (uncoupled) regime show typical diffraction-governed propagation for both fundamental and second-harmonic SPP beams (Fig. 8.12a, d). When the SPP intensity is gradually increased so that the nonlinear interaction becomes essential, the intensity distributions become deviating from the linear propagation regime. The observed intensity fringes are defined by the mismatch between the SPP phase velocities at the two frequencies. At the same time, the energy exchange between the beams can be seen: the maximum intensity of one beam corresponds to the minimum intensity of the other (Fig. 8.12c, f). Furthermore, the effect of narrowing of the SPP beams, driven by the nonlinearity, can be seen in the decrease in the average beam width for both the fundamental frequency and the second-harmonic frequency. The largest modulation of the intensity profiles is in the centres of the beams, where intensities are highest and the nonlinear coupling is strongest.

The efficiency of this soliton formation process can be increased if the phase matching between the fundamental and second-harmonic SPPs is achieved. This can be introduced using as the SPP-supporting interface a dielectric medium with anomalous dispersion to compensate for the SPP dispersion: $\text{Re}(n_1^{\text{eff}}) = \text{Re}(n_2^{\text{eff}})$ [1]. The evolution of the SPP beams from the linear to the nonlinear regime in this case shows the evident transformation of the SPP modes into highly localized nondiffracting solitons, allowing to achieve SPP beams with a narrower spatial profile than in the absence of the phase matching. SPP self-focusing effect occurring near the excitation boundary was also observed (Fig. 8.12i, l).

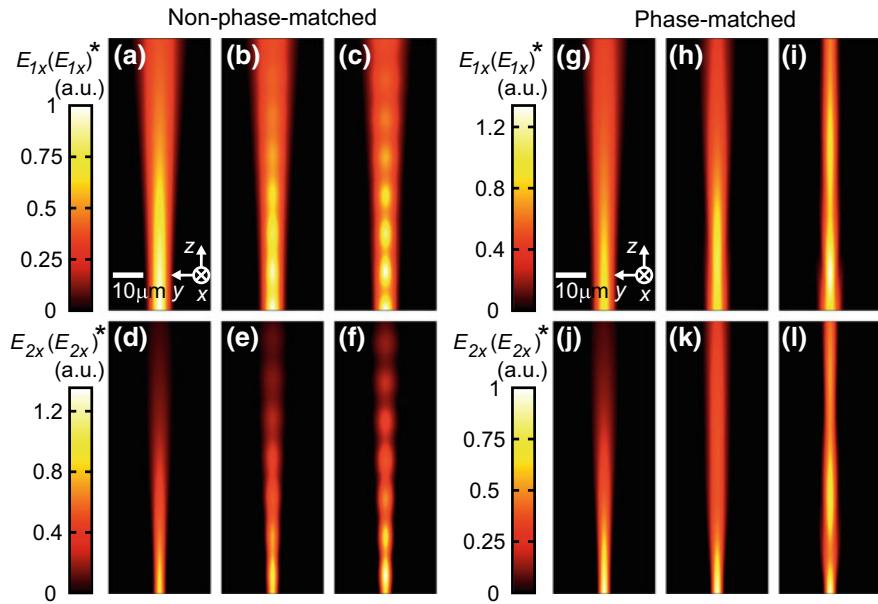


Fig. 8.12 **a–f** Evolution of co-propagating fundamental and SH SPP beams in the non-phase-matched case $2\beta_1 - \beta_2 \gg 0$. Linear propagation of **(a)** the fundamental and **(d)** the second-harmonic SPP beams in an effective 2D medium (Au/silica interface) with the effective indices $n_1^{\text{eff}} = 1.457 + 3.25 \times 10^{-4}i$ and $n_2^{\text{eff}} = 1.514 + 7.5 \times 10^{-4}i$ at 1500 and 750 nm wavelengths, respectively. Nonlinear propagation and self-focusing of **(b, c)** the fundamental and **(e, f)** the second-harmonic SPP beams for different light intensities corresponding to **(b, e)** $\chi^{(2)}E_1 = 0.02$ and **(c, g)** $\chi^{(2)}E_1 = 0.05$ nonlinearities. The initial amplitudes of both beams are equal ($E_1 = E_2$) and the beam half-widths are $w_{1,2} = 2.5\lambda_{1,2}$. **g–i** same as **a–f** but for the phase-matched case $2\beta_1 - \beta_2 = 0$ with $\text{Re}(n_1^{\text{eff}}) = \text{Re}(n_2^{\text{eff}}) = 1.583$. Adapted from [191]. Copyright 2013, Institute of Physics

8.7 Conclusion and Outlook

Various nanoplasmonic platforms based on planar and structured plasmonic-dielectric interfaces, plasmonic nanoparticles and their composites, as well as plasmonic metamaterials, facilitate enhancement and engineering of many nonlinear phenomena. Both spectrum and value of effective nonlinear susceptibilities can be manipulated in plasmonic environment. Nonlinear response can be introduced either by adjacent nonlinear optical materials in hybrid plasmonic structures or solely aided by inherent nonlinearities in plasmonic nanostructures themselves. In all cases, plasmonic structures provide strong field enhancement near the interfaces as well as flexible control over modal dispersions and strong sensitivity to the refractive index changes. All this enables plasmonic nanostructures to achieve both coherent and Kerr-type nonlinear responses on deep subwavelength scales, where the majority of conventional dielectric components fail due to natural limitations,

set by the diffraction limit and low values of nonlinearities requiring either long propagation range or very high light intensities. Flexible near-field manipulation in the vicinity of the nanostructures allows making use of spatially varying vectorial electromagnetic fields, and, thus, provides an opportunity to access all components of a nonlinear susceptibility tensor.

Another very important advantage of materials with free-electrons in the field of nonlinear optics is their inherent and ultrafast nonlinear response due to the complex dynamics of electrons in strong electromagnetic fields. The most widely used models for describing this nontrivial electron dynamics in nanostructures are the hydrodynamic model, taking into account free electron plasma with collective interactions between electrons, and the two-temperature model, taking into account non-equilibrium statistics of the carriers under optical excitation. Both models and respective nonlinear properties can be affected by geometrical constraints of nanostructured geometry, resulting in the importance of nonlocal electromagnetic effects, which become important at the length scale achievable with modern nanofabrication techniques. These nonlocal effects open a new route for engineering nonlinear response.

In this chapter, we gave a brief general overview of the field of nonlinear plasmonics. After discussing the origins of various nonlinear phenomena, plasmonic approaches for tailoring and enhancing nonlinearity were outlined, each time giving an illustration of related vivid physical phenomena. This included coherent nonlinear effects, such as harmonic generation at the nanoscale and nonlinear coupling of plasmonic resonances (an analogue of phase-matching at the macro-scale). A hydrodynamic time-domain numerical model was reviewed for coherent interactions of free-carrier plasma in the nanostructures of an arbitrary shape with an optical pulse of an arbitrary temporal profile, without any approximations. The approach allows to address the phenomena of multiple and resonantly-enhanced harmonic generation as well as reveals the interplay between the nonlocal effects and topology of a nanostructure. Ultrafast Kerr-type nonlinearity due to the saturation of interband transitions and heating of a free-electron gas in the conduction band was overviewed with a particular example of engineering magnitude, spectrum and sign of the nonlinearity using plasmonic nanorod metamaterial. At a designed wavelength, the metamaterial can provide either induced absorption or transparency and focusing or defocusing nonlinearity, depending on its geometrical parameters, opening up opportunities for custom engineered nonlinear materials for switching and modulation of light with light. Nonlinear surface waves were also briefly overviewed on the interfaces between linear and nonlinear media, including plasmon solitons and cascaded solitons.

The use of plasmonic nanostructures in nonlinear optics has allowed to move nonlinear optical processes in a realm of integrated photonics as well as to demonstrate low-light intensity nonlinear effects for free-standing optical applications. Switching, controlling, routing and manipulating of light with light is a difficult but rewarding problem with the applications ranging from on-chip optical data processing and routing to nonlinear optical components for optical communication networks and laser applications. Plasmonics has helped to demonstrate

many nonlinear optical effects on the subwavelength scales as well as in macroscopic implementations with much reduced required powers of light. Until recently, nonlinear plasmonic effects have mainly been demonstrated with conventional plasmonic materials, such as Au, Ag and Al. Unconventional phase changing plasmonic materials, such as Ga and VO₂ and ferromagnetic materials such as Co have been used in combination with plasmonic metals for controlling nonlinearity [192–195] as well as ferromagnetic materials (Ni) with plasmonic response in UV were also considered [196–198]. The development of a new type of plasmonic semiconductors based on nitrites and oxides, such as TiN, ITO, doped ZnO, as well as two-dimensional materials using plasmonic properties of graphene and transition metal dichalcogenides, including topological insulators and “Dirac-cone” materials, and plasmonic quantum dots has already started to provide new push in nonlinear plasmonics [199–202]. The advances in engineering of electric and magnetic resonances of high-refractive-index dielectric nanostructures, in particular with bulk second-order nonlinear susceptibility, have additionally provided new avenues for designing coherent nonlinearities in resonant structures [203]. The use of new features that become accessible with the new materials as well as a possibility to engineer, via appropriate nanostructuring in composite metamaterial arrangements, such features as near-zero permittivity and hyperbolic dispersion in a technologically relevant spectral ranges, additionally enhances practical applications of nonlinear plasmonic components. Modern fabrication technologies have permitted the achievement of fine nanoscale control over the plasmonic geometries which allowed flexible tuning of their resonances and succeeded in design of high effective nonlinearities at the desired wavelengths and with short response times. We believe that the future of nonlinear optics in its various facets of free-space, integrated or quantum nanophotonic technologies will be shaped by nonlinear nanostructures.

References

1. R.W. Boyd, *Nonlinear Optics* (Academic Press, 2008)
2. Y.R. Shen, *The Principles of Nonlinear Optics* (Wiley, 2002)
3. W.L. Barnes, A. Dereux, T.W. Ebbesen, Surface plasmon subwavelength optics. *Nature* **424**, 824–830 (2003)
4. A.V. Zayats, I.I. Smolyaninov, A.A. Maradudin, Nano-optics of surface plasmon polaritons. *Phys. Rep.* **408**, 131–314 (2005)
5. S.I. Bozhevolnyi (ed.), *Plasmonic Nanoguides and Circuits* (Pan Stanford Publishing Pte. Ltd., 2009)
6. A.V. Zayats, S. Maier (eds.), *Active Plasmonics and Tuneable Plasmonic Metamaterials* (Wiley, 2013)
7. W. Cai, V. Shalaev, *Optical Metamaterials: Fundamentals and Applications* (Springer, 2009)
8. A.A. Maradudin, J.R. Sambles, W.L. Barnes (eds.), *Modern Plasmonics* (Elsevier, 2014)
9. J.A. Schuller, E.S. Barnard, W.S. Cai, Y.C. Jun, J.S. White, M.L. Brongersma, Plasmonics for extreme light concentration and manipulation. *Nat. Mater.* **9**, 193–204 (2010)
10. M. Kauranen, A.V. Zayats, Nonlinear plasmonics. *Nat. Photonics* **6**, 737–748 (2012)

11. A.D. Boardman, A.V. Zayats, Nonlinear plasmonics, in *Modern Plasmonics*, ed. by A.A. Maradudin, J.R. Sambles, W.L. Barnes (Elsevier, 2014)
12. A.V. Krasavin, P. Ginzburg, A.V. Zayats, Free-electron optical nonlinearities in plasmonic nanostructures: a review of the hydrodynamic description. *Laser Photon. Rev.* **12**, 1700082 (2018)
13. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and Gratings* (Springer, 1988)
14. A.V. Krasavin, A.V. Zayats, Active nanophotonic circuitry based on dielectric-loaded plasmonic waveguides. *Adv. Opt. Mater.* **3**, 1662–1690 (2015)
15. S.M. Nie, S.R. Emery, Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. *Science* **275**, 1102–1106 (1997)
16. N.A. Mortensen, S. Raza, M. Wubs, T. Sondergaard, S.I. Bozhevolnyi, A generalized non-local optical response theory for plasmonic nanostructures. *Nat. Commun.* **5**, 3809 (2014)
17. P. Ginzburg, A.V. Zayats, Localized surface plasmon resonances in spatially dispersive nano-objects: phenomenological treatise. *ACS Nano* **7**, 4334–4342 (2013)
18. M. Agio, A. Alu (eds.), *Optical Antennas*, (Cambridge University Press, 2013)
19. A.G. Malshukov, Surface-enhanced Raman scattering. The present status. *Phys. Rep.-Rev. Sect. Phys. Lett.* **194**, 343–349 (1990)
20. J.D. Jackson, *Classical Electrodynamics* (Wiley, 1998)
21. C.M. Soukoulis, M. Wegener, Past achievements and future challenges in the development of three-dimensional photonic metamaterials. *Nat. Photonics* **5**, 523–530 (2011)
22. J.N. Anker, W.P. Hall, O. Lyandres, N.C. Shah, J. Zhao, R.P. Van Duyne, Biosensing with plasmonic nanosensors. *Nat. Mater.* **7**, 442–453 (2008)
23. A.V. Zayats, I.I. Smolyaninov, C.C. Davis, Observation of localized plasmonic excitations in thin metal films with near-field second-harmonic microscopy. *Opt. Commun.* **169**, 93–96 (1999)
24. F. Keilmann, R. Hillenbrand, Near-field microscopy by elastic light scattering from a tip. *Philos. Trans. R. Soc. London Ser. A* **362**, 787–805 (2004)
25. J. Butet, T.V. Raziman, K.Y. Yang, G.D. Bernasconi, O.J. Martin, Controlling the nonlinear optical properties of plasmonic nanoparticles with the phase of their linear response. *Opt. Express* **24**, 17138–17148 (2016)
26. G. Marino, P. Segovia, A.V. Krasavin, P. Ginzburg, N. Olivier, G.A. Wurtz, A.V. Zayats, Second-harmonic generation from hyperbolic plasmonic nanorod metamaterial slab. *Laser Photon. Rev.* **12**, 1700189 (2018)
27. P. Ginzburg, A. Krasavin, Y. Sonneneck, A. Murphy, R.J. Pollard, S.A. Maier, A.V. Zayats, Nonlinearly coupled localized plasmon resonances: resonant second-harmonic generation. *Phys. Rev. B* **86**, 085422 (2012)
28. A.V. Kabashin, P. Evans, S. Pastkovsky, W. Hendren, G.A. Wurtz, R. Atkinson, R. Pollard, V.A. Podolskiy, A.V. Zayats, Plasmonic nanorod metamaterials for biosensing. *Nat. Mater.* **8**, 867–871 (2009)
29. N. Vasilantonakis, G.A. Wurtz, V.A. Podolskiy, A.V. Zayats, Refractive index sensing with hyperbolic metamaterials: strategies for biosensing and nonlinearity enhancement. *Opt. Express* **23**, 14329–14343 (2015)
30. M.G. Banaee, K.B. Crozier, Mixed dimer double-resonance substrates for surface-enhanced raman spectroscopy. *ACS Nano* **5**, 307–314 (2011)
31. P. Genevet, J.P. Tetienne, E. Gatzogiannis, R. Blanchard, M.A. Kats, M.O. Scully, F. Capasso, Large enhancement of nonlinear optical phenomena by plasmonic nanocavity gratings. *Nano Lett.* **10**, 4880–4883 (2010)
32. M. Danckwerts, L. Novotny, Optical frequency mixing at coupled gold nanoparticles. *Phys. Rev. Lett.* **98**, 026104 (2007)
33. A.V. Krasavin, T.P. Vo, W. Dickson, P.M. Bolger, A.V. Zayats, All-plasmonic modulation via stimulated emission of copropagating surface plasmon polaritons on a substrate with gain. *Nano Lett.* **11**, 2231–2235 (2011)

34. A.V. Krasavin, S. Randhawa, J.S. Bouillard, J. Renger, R. Quidant, A.V. Zayats, Optically-programmable nonlinear photonic component for dielectric-loaded plasmonic circuitry. *Opt. Express* **19**, 25222–25229 (2011)
35. A.V. Krasavin, A.V. Zayats, All-optical active components for dielectric-loaded plasmonic waveguides. *Opt. Commun.* **283**, 1581–1584 (2010)
36. A.V. Krasavin, A.V. Zayats, Benchmarking system-level performance of passive and active plasmonic components: integrated circuit approach. *Proc. IEEE* **104**, 2338–2348 (2016)
37. K.F. MacDonald, Z.L. Samson, M.I. Stockman, N.I. Zheludev, Ultrafast active plasmonics. *Nat. Photonics* **3**, 55–58 (2009)
38. D. Pacifici, H.J. Lezec, H.A. Atwater, All-optical modulation by plasmonic excitation of CdSe quantum dots. *Nat. Photonics* **1**, 402–406 (2007)
39. R.A. Pala, K.T. Shimizu, N.A. Melosh, M.L. Brongersma, A nonvolatile plasmonic switch employing photochromic molecules. *Nano Lett.* **8**, 1506–1510 (2008)
40. R. Adato, H. Altug, In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas. *Nat. Commun.* **4**, 2154 (2013)
41. A.V. Krasavin, A.V. Zayats, Photonic signal processing on electronic scales: electro-optical field-effect nanoplasmonic modulator. *Phys. Rev. Lett.* **109**, 053901 (2012)
42. W.S. Cai, A.P. Vasudev, M.L. Brongersma, Electrically controlled nonlinear generation of light with plasmonics. *Science* **333**, 1720–1723 (2011)
43. I.-Y. Park, S. Kim, J. Choi, D.-H. Lee, Y.-J. Kim, M.F. Kling, M.I. Stockman, S.-W. Kim, Plasmonic generation of ultrashort extreme-ultraviolet light pulses. *Nat. Photonics* **5**, 677–681 (2011)
44. G.A. Wurtz, R. Pollard, W. Hendren, G.P. Wiederrecht, D.J. Gosztola, V.A. Podolskiy, A.V. Zayats, Designed ultrafast optical nonlinearity in a plasmonic nanorod metamaterial enhanced by nonlocality. *Nat. Nanotechnol.* **6**, 107–111 (2011)
45. A.D. Neira, N. Olivier, M.E. Nasir, W. Dickson, G.A. Wurtz, A.V. Zayats, Eliminating material constraints for nonlinearity with plasmonic metamaterials. *Nat. Commun.* **6**, 7757 (2015)
46. A.D. Neira, G.A. Wurtz, P. Ginzburg, A.V. Zayats, Ultrafast all-optical modulation with hyperbolic metamaterial integrated in Si photonic circuitry. *Opt. Express* **22**, 10987–10994 (2014)
47. G. Sartorello, N. Olivier, J.J. Zhang, W.S. Yue, D.J. Gosztola, G.P. Wiederrecht, G. Wurtz, A.V. Zayats, Ultrafast optical modulation of second- and third-harmonic generation from cut-disk-based metasurfaces. *ACS Photonics* **3**, 1517–1522 (2016)
48. R.W. Boyd, Z. Shi, I. De Leon, The third-order nonlinear optical susceptibility of gold. *Opt. Commun.* **326**, 74–79 (2014)
49. O. Lysenko, M. Bache, N. Olivier, A.V. Zayats, A. Lavrinenko, Nonlinear dynamics of ultrashort long-range surface plasmon polariton pulses in gold strip waveguides. *ACS Photonics* **3**, 2324–2329 (2016)
50. J. Dryzek, A. Czapla, Quantum size effect in optical spectra of thin metallic films. *Phys. Rev. Lett.* **58**, 721–724 (1987)
51. A.V. Zayats, O. Keller, K. Pedersen, A. Liu, F.A. Pudonin, Linear optical properties and second-harmonic generation from ultrathin niobium films: a search for quantization effects. *IEEE J. Quantum Electron.* **31**, 2044–2051 (1995)
52. N. Bloembergen, R.K. Chang, S.S. Jha, C.H. Lee, Optical second-harmonic generation in reflection from media with inversion symmetry. *Phys. Rev.* **174**, 813–822 (1968)
53. M. Perner, P. Bost, U. Lemmer, G. von Plessen, J. Feldmann, U. Becker, M. Mennig, M. Schmitt, H. Schmidt, Optically induced damping of the surface plasmon resonance in gold colloids. *Phys. Rev. Lett.* **78**, 2192–2195 (1997)
54. N.N. Lepeshkin, A. Schweinsberg, G. Piredda, R.S. Bennink, R.W. Boyd, Enhanced nonlinear optical response of one-dimensional metal-dielectric photonic crystals. *Phys. Rev. Lett.* **93**, 123902 (2004)

55. G. Piredda, D.D. Smith, B. Wendling, R.W. Boyd, Nonlinear optical properties of a gold-silica composite with high gold fill fraction and the sign change of its nonlinear absorption coefficient. *J. Opt. Soc. Am. B* **25**, 945–950 (2008)
56. M. Kono, M.M. Škorić, *Nonlinear Physics of Plasmas* (Springer, Berlin, Heidelberg, 2010)
57. P. Ginzburg, A.V. Krasavin, G.A. Wurtz, A.V. Zayats, Nonperturbative hydrodynamic model for multiple harmonics generation in metallic nanostructures. *ACS Photonics* **2**, 8–13 (2015)
58. A.V. Krasavin, P. Ginzburg, G.A. Wurtz, A.V. Zayats, Nonlocality-driven supercontinuum white light generation in plasmonic nanostructures. *Nat. Commun.* **7**, 11497 (2016)
59. M. Scalora, M.A. Vincenti, D. de Ceglia, V. Roppo, M. Centini, N. Akozbek, M.J. Bloember, Second- and third-harmonic generation in metal-based structures. *Phys. Rev. A* **82**, 043828 (2010)
60. D.W. Snoke, *Solid State Physics: Essential Concepts* (Addison-Wesley, 2009)
61. J.E. Sipe, V.C.Y. So, M. Fukui, G.I. Stegeman, Analysis of second-harmonic generation at metal surfaces. *Phys. Rev. B* **21**, 4389–4402 (1980)
62. P. Ginzburg, A. Hayat, N. Berkovitch, M. Orenstein, Nonlocal ponderomotive nonlinearity in plasmonics. *Opt. Lett.* **35**, 1551–1553 (2010)
63. R.H. Ritchie, Plasma losses by fast electrons in thin films. *Phys. Rev.* **106**, 874–881 (1957)
64. R. Sundaraman, P. Narang, A.S. Jermyn, W.A. Goddard III, H.A. Atwater, Theoretical predictions for hot-carrier generation from surface plasmon decay. *Nat. Commun.* **5**, 5788 (2014)
65. T. Higuchi, M.I. Stockman, P. Hommelhoff, Strong-field perspective on high-harmonic radiation from bulk solids. *Phys. Rev. Lett.* **113**, 213901 (2014)
66. N.W. Ashcroft, N. Mermin, *Solid State Physics*, (Brooks/Cole, 1976)
67. D.R. Nicholson, *Introduction to Plasma Theory* (Wiley, 1983)
68. L. Jiang, H.-L. Tsai, Improved two-temperature model and its application in ultrashort laser heating of metal films. *J. Heat Transfer* **127**, 1167–1173 (2005)
69. J. Bigot, J. Merle, O. Cregut, A. Daunois, Electron dynamics in copper metallic nanoparticles probed with femtosecond optical pulses. *Phys. Rev. Lett.* **75**, 4702–4705 (1995)
70. D. Pines, *Elementary Excitations in Solids: Lectures on Protons, Electrons, and Plasmons* (Perseus Books, 1999)
71. A. Marini, M. Conforti, G. Della Valle, H.W. Lee, T.X. Tran, W. Chang, M.A. Schmidt, S. Longhi, P.S.J. Russell, F. Biancalana, Ultrafast nonlinear dynamics of surface plasmon polaritons in gold nanowires due to the intrinsic nonlinearity of metals. *New J. Phys.* **15**, 013033 (2013)
72. S. Peruch, A. Neira, G.A. Wurtz, B. Wells, V.A. Podolskiy, A.V. Zayats, Geometry defines ultrafast hot carrier dynamics and Kerr nonlinearity in plasmonic metamaterial waveguides and cavities. *Adv. Opt. Mater.* **5**, 1700299 (2017)
73. J. Butet, P.F. Brevet, O.J.F. Martin, Optical second harmonic generation in plasmonic nanostructures: from fundamental principles to advanced applications. *ACS Nano* **9**, 10545–10562 (2015)
74. J. Rudnick, E.A. Stern, Second-harmonic radiation from metal surfaces. *Phys. Rev. B* **4**, 4274–4290 (1971)
75. D. Maystre, M. Neviere, R. Reinisch, Nonlinear polarisation inside metals: a mathematical study of the free-electron model. *Appl. Phys. A* **39**, 115–121 (1986)
76. X.M. Hua, J.I. Gersten, Theory of second-harmonic generation by small metal spheres. *Phys. Rev. B* **33**, 3756–3764 (1986)
77. D. Ostling, P. Stampfli, K.H. Bennemann, Theory of nonlinear optical properties of small metallic spheres. *Z. Phys. D* **28**, 169–175 (1993)
78. P. Guyot-Sionnest, Y.R. Shen, Bulk contribution in surface second-harmonic generation. *Phys. Rev. B* **38**, 7985–7989 (1988)

79. F.X. Wang, F.J. Rodríguez, W.M. Albers, R. Ahorinta, J.E. Sipe, M. Kauranen, Surface and bulk contributions to the second-order nonlinear optical response of a gold film. *Phys. Rev. B* **80**, 233402 (2009)
80. G. Bachelier, J. Butet, I. Russier-Antoine, C. Jonin, E. Benichou, P.F. Brevet, Origin of optical second-harmonic generation in spherical gold nanoparticles: local surface and nonlocal bulk contributions. *Phys. Rev. B* **82**, 235403 (2010)
81. A. Benedetti, M. Centini, C. Sibilia, M. Bertolotti, Engineering the second harmonic generation pattern from coupled gold nanowires. *J. Opt. Soc. Am. B* **27**, 408–416 (2010)
82. A. Benedetti, M. Centini, M. Bertolotti, C. Sibilia, Second harmonic generation from 3D nanoantennas: on the surface and bulk contributions by far-field pattern analysis. *Opt. Express* **19**, 26752–26767 (2011)
83. C. Forestiere, A. Capretti, G. Miano, Surface integral method for second harmonic generation in metal nanoparticles including both local-surface and nonlocal-bulk sources. *J. Opt. Soc. Am. B* **30**, 2355–2364 (2013)
84. A. Capretti, C. Forestiere, L. Dal Negro, G. Miano, Full-wave analytical solution of second-harmonic generation in metal nanospheres. *Plasmonics* **9**, 151–166 (2013)
85. J.I. Dadap, J. Shan, K.B. Eisenthal, T.F. Heinz, Second-harmonic Rayleigh scattering from a sphere of centrosymmetric material. *Phys. Rev. Lett.* **83**, 4045–4048 (1999)
86. J.I. Dadap, J. Shan, T.F. Heinz, Theory of optical second-harmonic generation from a sphere of centrosymmetric material: small-particle limit. *J. Opt. Soc. Am. B* **21**, 1328–1347 (2004)
87. Y. Pavlyukh, W. Hübner, Nonlinear Mie scattering from spherical particles. *Phys. Rev. B* **70**, 245434 (2004)
88. J. Nappa, G. Revillo, I. Russier-Antoine, E. Benichou, C. Jonin, P.F. Brevet, Electric dipole origin of the second harmonic generation of small metallic particles. *Phys. Rev. B* **71**, 165407 (2005)
89. I. Russier-Antoine, E. Benichou, G. Bachelier, C. Jonin, P.F. Brevet, Multipolar contributions of the second harmonic generation from silver and gold nanoparticles. *J. Phys. Chem. C* **111**, 9044–9048 (2007)
90. G. Bachelier, I. Russier-Antoine, E. Benichou, C. Jonin, P.F. Brevet, Multipolar second-harmonic generation in noble metal nanoparticles. *J. Opt. Soc. Am. B* **25**, 955–960 (2008)
91. A.V. Smolyaninov II, C.C. Zayats, Davis: near-field second harmonic generation from a rough metal surface. *Phys. Rev. B* **56**, 9290–9293 (1997)
92. A.V. Zayats, T. Kalkbrenner, V. Sandoghdar, J. Mlynek, Second-harmonic generation from individual surface defects under local excitation. *Phys. Rev. B* **61**, 4545–4548 (2000)
93. R. Kolkowski, J. Szeszko, B. Dwir, E. Kapon, J. Zyss, Non-centrosymmetric plasmonic crystals for second-harmonic generation with controlled anisotropy and enhancement. *Laser Photon. Rev.* **10**, 287–298 (2016)
94. B.L. Wang, R. Wang, R.J. Liu, X.H. Lu, J. Zhao, Z.Y. Li, Origin of shape resonance in second-harmonic generation from metallic nanohole arrays. *Sci. Rep.* **3**, 2358 (2013)
95. K. O'Brien, H. Suchowski, J. Rho, A. Salandrino, B. Kante, X. Yin, X. Zhang, Predicting nonlinear properties of metamaterials from the linear response. *Nat. Mater.* **14**, 379–383 (2015)
96. M. Celebrano, X. Wu, M. Baselli, S. Grossmann, P. Biagioni, A. Locatelli, C. De Angelis, G. Cerullo, R. Osellame, B. Hecht, L. Duo, F. Ciccacci, M. Finazzi, Mode matching in multiresonant plasmonic nanoantennas for enhanced second harmonic generation. *Nat. Nanotechnol.* **10**, 412–417 (2015)
97. M.L. Ren, S.Y. Liu, B.L. Wang, B.Q. Chen, J. Li, Z.Y. Li, Giant enhancement of second harmonic generation by engineering double plasmonic resonances at nanoscale. *Opt. Express* **22**, 28653–28661 (2014)
98. B.K. Canfield, H. Husu, J. Laukkonen, B.F. Bai, M. Kuittinen, J. Turunen, M. Kauranen, Local field asymmetry drives second-harmonic generation in noncentrosymmetric nanodimers. *Nano Lett.* **7**, 1251–1255 (2007)

99. B. Wells, A.Y. Bykov, G. Marino, M.E. Nasir, A.V. Zayats, V.A. Podolskiy, Structural second-order nonlinearity in plasmonic metamaterials. *Optica* **5**, 1502–1507 (2018)
100. M.I. Shalaev, Z.A. Kudyshev, N.M. Litchinitser, Twisted light in a nonlinear mirror. *Opt. Lett.* **38**, 4288–4291 (2013)
101. K.A. O'Donnell, R. Torre, C.S. West, Observations of second-harmonic generation from randomly rough metal surfaces. *Phys. Rev. B* **55**, 7985–7992 (1997)
102. M.I. Stockman, D.J. Bergman, C. Ancea, S. Brasselet, J. Zyss, Enhanced second-harmonic generation by metal surfaces with nanoscale roughness: nanoscale dephasing, depolarization, and correlations. *Phys. Rev. Lett.* **92**, 057402 (2004)
103. T. Stefaniuk, N. Olivier, A. Belardini, C.P.T. McPolin, C. Sibilia, A.A. Wronkowska, A. Wronkowski, T. Szoplik, A.V. Zayats, Self-assembled silver-germanium nanolayer metamaterial with the enhanced nonlinear response. *Adv. Opt. Mater.* **5**, 1700753 (2017)
104. J.P. Dewitz, W. Hubner, K.H. Bennemann, Theory for nonlinear Mie-scattering from spherical metal clusters. *Z. Phys. D* **37**, 75–84 (1996)
105. D. Carroll, X.H. Zheng, Spatial and angular distributions of third harmonic generation from metal surfaces. *Eur. Phys. J. D* **5**, 135–144 (1999)
106. Y. Yu, S.-S. Fan, H.-W. Dai, Z.-W. Ma, X. Wang, J.-B. Han, L. Li, Plasmon resonance enhanced large third-order optical nonlinearity and ultrafast optical response in Au nanobipyramids. *Appl. Phys. Lett.* **105**, 061903 (2014)
107. B. Metzger, M. Hentschel, M. Nesterov, T. Schumacher, M. Lippitz, H. Giessen, Nonlinear optics of complex plasmonic structures: linear and third-order optical response of orthogonally coupled metallic nanoantennas. *Appl. Phys. B* **122**, 77 (2016)
108. J.B. Lassiter, X. Chen, X. Liu, C. Ciraci, T.B. Hoang, S. Larouche, S.-H. Oh, M.H. Mikkelsen, D.R. Smith, Third-harmonic generation enhancement by film-coupled plasmonic stripe resonators. *ACS Photonics* **1**, 1212–1217 (2014)
109. K. Li, X. Li, D. Yuan Lei, S. Wu, Y. Zhan, Plasmon gap mode-assisted third-harmonic generation from metal film-coupled nanowires. *Appl. Phys. Lett.* **104**, 261105 (2014)
110. M.S. Nezami, R. Gordon, Localized and propagating surface plasmon resonances in aperture-based third harmonic generation. *Opt. Express* **23**, 32006–32014 (2015)
111. G. Hajisalem, D.K. Hore, R. Gordon, Interband transition enhanced third harmonic generation from nanoplasmonic gold. *Opt. Mater. Express* **5**, 2217–2224 (2015)
112. T. Wu, P.P. Shum, Y. Sun, X. Shao, T. Huang, Study on the crucial conditions for efficient third harmonic generation using a metal-hybrid-metal plasmonic slot waveguide. *Opt. Express* **23**, 253–263 (2015)
113. I.D. Mayergoyz, D.R. Fredkin, Z. Zhang, Electrostatic (plasmon) resonances in nanoparticles. *Phys. Rev. B* **72**, 155412 (2005)
114. J. Butet, G. Bachelier, I. Russier-Antoine, C. Jonin, E. Benichou, P.F. Brevet, Interference between selected dipoles and octupoles in the optical second-harmonic generation from spherical gold nanoparticles. *Phys. Rev. Lett.* **105**, 077401 (2010)
115. D.V. Guzatov, V.V. Klimov, M.Y. Pikhota, Plasmon oscillations in ellipsoid nanoparticles: beyond dipole approximation. *Laser Phys.* **20**, 85–99 (2009)
116. S. Asano, G. Yamamoto, Light-scattering by a spheroidal particle. *Appl. Opt.* **14**, 29–49 (1975)
117. P. Segovia, G. Marino, A.V. Krasavin, N. Olivier, G.A. Wurtz, P.A. Belov, P. Ginzburg, A. V. Zayats, Hyperbolic metamaterial antenna for second-harmonic generation tomography. *Opt. Express* **23**, 30730–30738 (2015)
118. E.V. Makeev, S.E. Skipetrov, Second harmonic generation in suspensions of spherical particles. *Opt. Commun.* **224**, 139–147 (2003)
119. J.I. Dadap, Optical second-harmonic scattering from cylindrical particles. *Phys. Rev. B* **78**, 205322 (2008)
120. S. Varró, F. Ehlotzky, Higher-harmonic generation from a metal surface in a powerful laser field. *Phys. Rev. A* **49**, 3106–3109 (1994)

121. G. Farkas, C. Tóth, S.D. Moustazis, N.A. Papadogiannis, C. Fotakis, Observation of multiple-harmonic radiation induced from a gold surface by picosecond neodymium-doped yttrium aluminum garnet laser pulses. *Phys. Rev. A* **46**, R3605–R3608 (1992)
122. C.F. Bohren, D.R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley-Interscience, New York, 1983)
123. S. Raza, S.I. Bozhevolnyi, M. Wubs, N. Asger Mortensen, Nonlocal optical response in metallic nanostructures. *J. Phys. Condens. Matter* **27**, 183204 (2015)
124. G. Toscano, J. Straubel, A. Kwiatkowski, C. Rockstuhl, F. Evers, H. Xu, N.A. Mortensen, M. Wubs, Resonance shifts and spill-out effects in self-consistent hydrodynamic nanoplasmonics. *Nat. Commun.* **6**, 7132 (2015)
125. R.B. Davidson II, J.I. Ziegler, G. Vargas, S.M. Avanesyan, Y. Gong, W. Hess, R.F. Haglund Jr, Efficient forward second-harmonic generation from planar archimedean nanospirals. *Nanophotonics* **4**, 108–113 (2015)
126. H.M. Gibbs: *Optical Bistability: Controlling Light with Light* (Academic Press, 1985)
127. C.K. Sun, F. Vallée, L. Acioli, E.P. Ippen, J.G. Fujimoto, Femtosecond investigation of electron thermalization in gold. *Phys. Rev. B* **48**, 12365–12368 (1993)
128. N. Del Fatti, R. Bouffanais, F. Vallee, C. Flytzanis, Nonequilibrium electron interactions in metal films. *Phys. Rev. Lett.* **81**, 922–925 (1998)
129. M.I. Stockman, Nanoplasmonics: past, present, and glimpse into future. *Opt. Express* **19**, 22029–22106 (2011)
130. W. Dickson, G.A. Wurtz, P.R. Evans, R.J. Pollard, A.V. Zayats, Electronically controlled surface plasmon dispersion and optical transmission through metallic hole arrays using liquid crystal. *Nano Lett.* **8**, 281–286 (2008)
131. R.A. Innes, J.R. Sambles, Optical non-linearity in liquid crystals using surface plasmon-polaritons. *J. Phys. Condens. Matter* **1**, 6231–6260, 021 (1989)
132. *Optical Properties of Organic Molecules and Crystals* (Academic Press, 1987)
133. S. Link, M.A. El-Sayed, Spectral properties and relaxation dynamics of surface plasmon electronic oscillations in gold and silver nanodots and nanorods. *J. Phys. Chem. B* **103**, 8410–8426 (1999)
134. M. Pelton, J. Aizpurua, G. Bryant, Metal-nanoparticle plasmonics. *Laser Photon. Rev.* **2**, 136–159 (2008)
135. H. Baida, D. Mongin, D. Christofilos, G. Bachelier, A. Crut, P. Maioli, N. Del Fatti, F. Vallee, Ultrafast nonlinear optical response of a single gold nanorod near its surface plasmon resonance. *Phys. Rev. Lett.* **107**, 057402 (2011)
136. M. Halonen, A.A. Lipovskii, Y.P. Svirko, Femtosecond absorption dynamics in glass-metal nanopcomposites. *Opt. Express* **15**, 6840–6845 (2007)
137. G. Ma, W. Sun, S.H. Tang, H. Zhang, Z. Shen, S. Qian, Size and-dielectric dependence of the third-order nonlinear optical response of Au nanocrystals embedded in matrices. *Opt. Lett.* **27**, 1043–1045 (2002)
138. W. Dickson, G.A. Wurtz, P. Evans, D. O'Connor, R. Atkinson, R. Pollard, A.V. Zayats, Dielectric-loaded plasmonic nanoantenna arrays: a metamaterial with tuneable optical properties. *Phys. Rev. B* **76**, 115411 (2007)
139. M. Fu, K. Wang, H. Long, G. Yang, P. Lu, F. Hetsch, A.S. Susha, A.L. Rogach, Resonantly enhanced optical nonlinearity in hybrid semiconductor quantum dot—metal nanoparticle structures. *Appl. Phys. Lett.* **100**, 063117 (2012)
140. M. Abb, P. Albella, J. Aizpurua, O.L. Muskens, All-optical control of a single plasmonic nanoantenna-ITO hybrid. *Nano Lett.* **11**, 2457–2463 (2011)
141. I.S. Maksymov, A.E. Miroshnichenko, Y.S. Kivshar, Actively tunable bistable optical Yagi-Uda nanoantenna. *Opt. Express* **20**, 8929–8938 (2012)
142. I.I. Smolyaninov, Quantum fluctuations of the refractive index near the interface between a metal and a nonlinear dielectric. *Phys. Rev. Lett.* **94**, 057403 (2005)
143. I. Smolyaninov, A.V. Zayats, A. Gungor, C.C. Davis, Single-photon tunneling via localized surface plasmons. *Phys. Rev. Lett.* **88**, 187402 (2002)

144. I.I. Smolyaninov, C.C. Davis, A.V. Zayats, Light-controlled photon tunneling. *Appl. Phys. Lett.* **81**, 3314–3316 (2002)
145. Y. Lin, X. Zhang, X. Fang, S. Liang, A cross-stacked plasmonic nanowire network for high-contrast femtosecond optical switching. *Nanoscale* **8**, 1421–1429 (2016)
146. G. Della Valle, D. Polli, P. Biagioni, C. Martella, M.C. Giordano, M. Finazzi, S. Longhi, L. Duò, G. Cerullo, F. Bautier de Mongeot, Self-organized plasmonic metasurfaces for all-optical modulation. *Phys. Rev. B* **91**, 235440 (2015)
147. M. Pohl, V.I. Belotelov, I.A. Akimov, S. Kasture, A.S. Vengurlekar, A.V. Gopal, A.K. Zvezdin, D.R. Yakovlev, M. Bayer, Plasmonic crystals for ultrafast nanophotonics: optical switching of surface plasmon polaritons. *Phys. Rev. B* **85**, 081401(R) (2012)
148. X. Wang, R. Morea, J. Gonzalo, B. Palpant, Coupling localized plasmonic and photonic modes tailors and boosts ultrafast light modulation by gold nanoparticles. *Nano Lett.* **15**, 2633–2639 (2015)
149. C.P.T. McPolin, N. Olivier, J.-S. Bouillard, D. O'Connor, A.V. Krasavin, W. Dickson, G.A. Wurtz, A.V. Zayats, Universal switching of plasmonic signals using optical resonator modes. *Light Sci. Appl.* **6**, e16237 (2017)
150. L.H. Nicholls, F.J. Rodriguez-Fortuno, M.E. Nasir, R.M. Cordova-Castro, N. Olivier, G.A. Wurtz, A.V. Zayats, Ultrafast synthesis and switching of light polarization in nonlinear anisotropic metamaterials. *Nat. Photonics* **11**, 628–633 (2017)
151. N.E. Khokhlov, D.O. Ignatyeva, V.I. Belotelov, Plasmonic pulse shaping and velocity control via photoexcitation of electrons in a gold film. *Opt. Express* **22**, 28019–28026 (2014)
152. A.V. Krasavin, N.I. Zheludev, Active plasmonics: controlling signals in Au/Ga waveguide using nanoscale structural transformations. *Appl. Phys. Lett.* **84**, 1416–1418 (2004)
153. A.V. Krasavin, K.F. MacDonald, N.I. Zheludev, A.V. Zayats, High-contrast modulation of light with light by control of surface plasmon polariton wave coupling. *Appl. Phys. Lett.* **85**, 3369–3371 (2004)
154. A.V. Krasavin, A.V. Zayats, N.I. Zheludev, Active control of surface plasmon–polariton waves. *J. Opt. A Pure Appl. Opt.* **7**, S85–S89 (2005)
155. N. Rotenberg, M. Betz, H.M. van Driel, Ultrafast all-optical coupling of light to surface plasmon polaritons on plain metal surfaces. *Phys. Rev. Lett.* **105**, 017402 (2010)
156. A.V. Krasavin, A.V. Zayats, Three-dimensional numerical modeling of photonic integration with dielectric-loaded SPP waveguides. *Phys. Rev. B* **78**, 045425 (2008)
157. S. Randhawa, A.V. Krasavin, T. Holmgård, J. Renger, S.I. Bozhevolnyi, A.V. Zayats, R. Quidant, Experimental demonstration of dielectric-loaded plasmonic waveguide disk resonators at telecom wavelengths. *Appl. Phys. Lett.* **98**, 161102 (2011)
158. G.A. Wurtz, A.V. Zayats, Nonlinear surface plasmon polaritonic crystals. *Laser Photonics Rev.* **2**, 125–135 (2008)
159. V. Mikhailov, G.A. Wurtz, J. Elliott, P. Bayvel, A.V. Zayats, Dispersing light with surface plasmon polaritonic crystals. *Phys. Rev. Lett.* **99**, 083901 (2007)
160. A. Minovich, J. Farnell, D.N. Neshev, I. McKerracher, F. Karouta, J. Tian, D.A. Powell, I.V. Shadrivov, H. Hoe Tan, C. Jagadish, Y.S. Kivshar, Liquid crystal based nonlinear fishnet metamaterials. *Appl. Phys. Lett.* **100**, 121113 (2012)
161. I.I. Smolyaninov, A.V. Zayats, A. Stinishevsky, C.C. Davis, Optical control of photon tunneling through an array of nanometer-scale cylindrical channels. *Phys. Rev. B* **66**, 205414 (2002)
162. G.A. Wurtz, R. Pollard, A.V. Zayats, Optical bistability in nonlinear surface-plasmon polaritonic crystals. *Phys. Rev. Lett.* **97**, 057402 (2006)
163. N. Rotenberg, M. Betz, H.M. Van Driel, Ultrafast control of grating-assisted light coupling to surface plasmons. *Opt. Lett.* **33**, 2137–2139 (2008)
164. K.T. Tsai, G.A. Wurtz, J.Y. Chu, T.Y. Cheng, H.H. Wang, A.V. Krasavin, J.H. He, B.M. Wells, V.A. Podolskiy, J.K. Wang, Y.L. Wang, A.V. Zayats, Looking into meta-atoms of plasmonic nanowire metamaterial. *Nano Lett.* **14**, 4971–4976 (2014)

165. P. Ginzburg, D.J. Roth, M.E. Nasir, P. Segovia, A.V. Krasavin, J. Levitt, L.M. Hirvonen, B. Wells, K. Suhling, D. Richards, V.A. Podolskiy, A.V. Zayats, Spontaneous emission in non-local materials. *Light Sci. Appl.* **6**, e16273 (2017)
166. R.J. Pollard, A. Murphy, W.R. Hendren, P.R. Evans, R. Atkinson, G.A. Wurtz, A.V. Zayats, V.A. Podolskiy, Optical nonlocalities and additional waves in epsilon-near-zero metamaterials. *Phys. Rev. Lett.* **102**, 127405 (2009)
167. B.M. Wells, A.V. Zayats, V.A. Podolskiy, Nonlocal optics of plasmonic nanowire metamaterials. *Phys. Rev. B* **89**, 10, 035111 (2014)
168. V.A. Podolskiy, P. Ginzburg, B. Wells, A.V. Zayats, Light emission in nonlocal plasmonic metamaterials. *Faraday Discuss.* **178**, 61–70 (2015)
169. A.E. Nikolaenko, F. De Angelis, S.A. Boden, N. Papasimakis, P. Ashburn, E. Di Fabrizio, N.I. Zheludev, Carbon nanotubes in a photonic metamaterial. *Phys. Rev. Lett.* **104**, 4, 153902 (2010)
170. M. Ren, B. Jia, J.Y. Ou, E. Plum, J. Zhang, K.F. MacDonald, A.E. Nikolaenko, J. Xu, M. Gu, N.I. Zheludev, Nanostructured plasmonic medium for terahertz bandwidth all-optical switching. *Adv. Mater.* **23**, 5540–5544 (2011)
171. G.A. Wurtz, P.R. Evans, W. Hendren, R. Atkinson, W. Dickson, R.J. Pollard, A.V. Zayats, W. Harrison, C. Bower, Molecular plasmonics with tunable exciton-plasmon coupling strength in J-aggregate hybridized Au nanorod assemblies. *Nano Lett.* **7**, 1297–1303 (2007)
172. N. Vasilantonakis, M.E. Nasir, W. Dickson, G.A. Wurtz, A.V. Zayats, Bulk plasmon-polaritons in hyperbolic nanorod metamaterial waveguides. *Laser Photon. Rev.* **9**, 345–353 (2015)
173. M.Z. Alam, I. De Leon, R.W. Boyd, Large optical nonlinearity of indium tin oxide in its epsilon-near-zero region. *Science* **352**, 795–797 (2016)
174. D. Mihalache, M. Bertolotti, C. Sibilia, Nonlinear wave propagation in planar structures, ed. by E. Wolf in *Progress in Optics* vol 27 (Elsevier, 1989)
175. P. Ginzburg, E. Hirshberg, M. Orenstein, Rigorous analysis of vectorial plasmonic diffraction: single- and double-slit experiments. *J. Opt. A Pure Appl. Opt.* **11**, 114024 (2009)
176. D.K. Gramotnev, S.I. Bozhevolnyi, Plasmonics beyond the diffraction limit. *Nat. Photonics* **4**, 83–91 (2010)
177. P. Ginzburg, D. Arbel, M. Orenstein, Gap plasmon polariton structure for very efficient microscale-to-nanoscale interfacing. *Opt. Lett.* **31**, 3288–3290 (2006)
178. P. Ginzburg, M. Orenstein, Plasmonic transmission lines: from micro to nano scale with lambda/4 impedance matching. *Opt. Express* **15**, 6762–6767 (2007)
179. I.D. Rukhlenko, A. Pannipitiya, M. Premaratne, G.P. Agrawal, Exact dispersion relation for nonlinear plasmonic waveguides. *Phys. Rev. B* **84**, 113409 (2011)
180. P. Ginzburg, M. Orenstein, Nonlinear effects in plasmonic systems, in *Active Plasmonics and Tuneable Plasmonic Metamaterials*, ed. by A.V. Zayats, S.A. Maier (Wiley, 2013)
181. E. Feigenbaum, M. Orenstein, Plasmon-soliton. *Opt. Lett.* **32**, 674–676 (2007)
182. E. Feigenbaum, M. Orenstein, Modeling of complementary (void) plasmon waveguiding. *J. Lightwave Technol.* **25**, 2547–2562 (2007)
183. A.R. Davoyan, I.V. Shadrivov, Y.S. Kivshar, Self-focusing and spatial plasmon-polariton solitons. *Opt. Express* **17**, 21732–21737 (2009)
184. A.R. Davoyan, I.V. Shadrivov, A.A. Zharov, D.K. Gramotnev, Y.S. Kivshar, Nonlinear nanofocusing in tapered plasmonic waveguides. *Phys. Rev. Lett.* **105**, 116804 (2010)
185. A. Marini, D.V. Skryabin, B. Malomed, Stable spatial plasmon solitons in a dielectric-metal-dielectric geometry with gain and loss. *Opt. Express* **19**, 6616–6622 (2011)
186. Y. Liu, G. Bartal, D.A. Genov, X. Zhang, Subwavelength discrete solitons in nonlinear metamaterials. *Phys. Rev. Lett.* **99**, 153901 (2007)
187. Y.N. Karamzin, A.P. Sukhorukov, Mutual focusing of high-power light beams in media with quadratic nonlinearity. *Sov. Phys. JETP* **41**, 414–420 (1976)
188. W.E. Torruellas, Z. Wang, D.J. Hagan, E.W. VanStryland, G.I. Stegeman, L. Torner, C.R. Menyuk, Observation of two-dimensional spatial solitary waves in a quadratic medium. *Phys. Rev. Lett.* **74**, 5036–5039 (1995)

189. R. Schiek, Y. Baek, G.I. Stegeman, One-dimensional spatial solitary waves due to cascaded second-order nonlinearities in planar waveguides. *Phys. Rev. E* **53**, 1138–1141 (1996)
190. L. Torner, A. Barthelemy, Quadratic solitons: recent developments. *IEEE J. Quantum Electron.* **39**, 22–30 (2003)
191. P. Ginzburg, A.V. Krasavin, A.V. Zayats, Cascaded second-order surface plasmon solitons due to intrinsic metal nonlinearity. *New J. Phys.* **15**, 013031 (2013)
192. A.V. Krasavin, K.F. MacDonald, A.S. Schwanecke, N.I. Zheludev, Gallium/aluminum nanocomposite material for nonlinear optics and nonlinear plasmonics. *Appl. Phys. Lett.* **89**, 031118 (2006)
193. T.V. Murzina, T.V. Misuryaev, A.F. Kravets, J. Gudde, D. Schuhmacher, G. Marowsky, A. A. Nikulin, O.A. Aktsipetrov, Nonlinear magneto-optical Kerr effect and plasmon-assisted SHG in magnetic nanomaterials exhibiting giant magnetoresistance. *Surf. Sci.* **482**, 1101–1106 (2001)
194. I. Razdolski, D. Makarov, O.G. Schmidt, A. Kirilyuk, T. Rasing, V.V. Temnov, Nonlinear surface magnetoplasmonics in Kretschmann multilayers. *ACS Photonics* **3**, 179–183 (2016)
195. T. Jostmeier, M. Mangold, J. Zimmer, H. Karl, H.J. Krenner, C. Ruppert, M. Betz, Thermochromic modulation of surface plasmon polaritons in vanadium dioxide nanocomposites. *Opt. Express* **24**, 17321–17331 (2016)
196. V.L. Krutynskiy, I.A. Kolmychek, E.A. Gan'shina, T.V. Murzina, P. Evans, R. Pollard, A. A. Stashkevich, G.A. Wurtz, A.V. Zayats, Plasmonic enhancement of nonlinear magneto-optical response in nickel nanorod metamaterials. *Phys. Rev. B* **87**, 035116 (2013)
197. V.K. Valev, A.V. Silhanek, W. Gillijns, Y. Jeyaram, H. Paddubrouskaya, A. Volodin, C.G. Biris, N.C. Panouli, B. De Clercq, M. Ameloot, O.A. Aktsipetrov, V.V. Moshchalkov, T. Verbiest, Plasmons reveal the direction of magnetization in nickel nanostructures. *ACS Nano* **5**, 91–96 (2011)
198. V. Bonanni, S. Bonetti, T. Pakizeh, Z. Pirzadeh, J.N. Chen, J. Nogues, P. Vavassori, R. Hillenbrand, J. Akerman, A. Dmitriev, Designer magnetoplasmonics with nickel nanoferromagnets. *Nano Lett.* **11**, 5333–5338 (2011)
199. A. Boltasseva, H.A. Atwater, Low-loss plasmonic metamaterials. *Science* **331**, 290–291 (2011)
200. H. Zhang, S. Virally, Q.L. Bao, L.K. Ping, S. Massar, N. Godbout, P. Kockaert, Z-scan measurement of the nonlinear refractive index of graphene. *Opt. Lett.* **37**, 1856–1858 (2012)
201. J.D. Cox, I. Silveiro, F.J.G. de Abajo, Quantum effects in the nonlinear response of graphene plasmons. *ACS Nano* **10**, 1995–2003 (2016)
202. R.I. Woodward, R.T. Murray, C.F. Phelan, R.E.P. de Oliveira, T.H. Runcorn, E.J.R. Kelleher, S. Li, E.C. de Oliveira, G.J.M. Fechine, G. Eda, C.J.S. de Matos, Characterization of the second- and third-order nonlinear optical susceptibilities of monolayer MoS₂ using multiphoton microscopy. *2D Mater.* **4**, 011006 (2016)
203. M. Rahmani, G. Leo, I. Brener, A.V. Zayats, S. Maier, C. De Angelis, H. Tan, V.F. Gili, F. Karouta, R. Oulton, K. Vora, M. Lysevych, I. Staude, L. Xu, A. Miroshnichenko, C. Jagadish, D. Neshev, Nonlinear frequency conversion in optical nanoantennas and metasurfaces: materials evolution and fabrication. *Opto-Electron. Adv.* **1**, 180021 (2018)

Chapter 9

Nonlinear Optics with Elliptically Polarized Singular Beams and Short Pulses in Media with Spatial Dispersion



Vladimir A. Makarov

Abstract The conditions of appearance and the behavior of polarization singularities (*C*-points) in the cross-section of light beam arising due to nonlinear interaction of elliptically polarized laser beams with a medium with nonlocality of quadratic and cubic optical responses are discussed. The formation dynamics and propagation features of *C*-points, including pairwise creation and annihilation, for sum-frequency and second harmonic generation, beams self-action and interaction and other nonlinear optical processes are presented. We also discuss the effects accompanying the propagation of ultrashort (several oscillations) elliptically polarized light pulses in nonlinear isotropic gyrotrropic medium with frequency dispersion.

9.1 Introduction

A high-speed development of nonlinear optics, which opened to the world a variety of unusually beautiful physical phenomena, first overshadowed for some time the analysis of changes in polarizations of beams and pulses interacting in a medium (the history of the discovery and investigations of main nonlinear optical effects is presented in papers [1–5]). It was assumed initially that the states of their polarizations can produce only an insignificant influence on classical effects of nonlinear optics and, hence, a rather time consuming theoretical study involving the solution of at least twice as many coupled nonlinear partial differential equations than in the approximation that the polarization of light remains constant during light propagation, is hardly justified and is of academic interest only. Moreover, such studies were not stimulated experimentally. The estimates showed that some polarization effects could be observed only at laser radiation intensities quite large for mid-1960s.

V. A. Makarov (✉)

Faculty of Physics and International Laser Center, Lomonosov Moscow State University, Moscow, Russia

e-mail: vamakarov@phys.msu.ru; makarov@ilc.edu.ru

In 1950, Sergey Vavilov predicted that birefringence, dichroism and polarization rotatory power depend on light intensity [6]. The self-rotation of the polarization ellipse, which became stronger with the increase of the polarization ellipse ellipticity degree in the incident plane wave and completely disappears for the linearly polarized light, was observed by Terhune [7]. S.A. Akhmanov and V.I. Zharikov predicted in 1967 the effect of nonlinear optical activity [8]—intensity-dependent rotation of the polarization plane of linearly polarized light falling on a medium with a spatial dispersion of the cubic nonlinearity. These investigations stimulated subsequent development of nonlinear polarized optics. Theoretical and experimental studies performed to date make it possible to assert definitely that the effects of polarization self-action and interaction of waves belong to delicate but widespread effects of nonlinear optics. The use of the approximation of polarization wave invariance during the propagation in theoretical calculations is hardly justified and represents only the first step to the consequent description of nonlinear optical phenomena. A wave in quantum electronic devices is always elliptically polarized, the degree of its ellipticity and the inclination angle of the principal axis of the polarization ellipse changing during the propagation through nonlinear crystals because of reflections from smooth surfaces and also because of resonator effects. Furthermore, when waves interact in nonlinear media, their polarization can change differently at different points of the light beam cross section [9–13]. In a number of cases, an elliptically polarized pulse can split into separate parts, the modulus of the degree of the electric field ellipticity in each part being close to unity. In this case, the rotation direction of the electric field vector in the pulse center is opposite to the rotation direction in side parts.

The extensive list of spectroscopic schemes involves methods based on the use of intensity-dependent variations in light polarization. Being one of the most advanced, the method for polarization measurements [14] allows one to detect rather weak changes in the degree of ellipticity and the rotation angle of the principal axis of the polarization ellipse of the signal wave and, hence, to obtain important spectroscopic data on the matter, which are unavailable with the help of other investigation techniques. Additional possibilities are related to the use of elliptically polarized fundamental waves in optical experiments. The latter makes it possible to increase the number of “degrees of freedom”. In other words, one can change the degree of ellipticity, the mutual orientation of the principal axes of polarization ellipses and other parameters of the fundamental waves and thereby to emphasize or suppress the contribution of local, nonlocal and surface nonlinear susceptibilities to the intensity and polarization of a signal wave.

Despite the popularity and wide range of the considered problems of singular linear polarization optics, investigations of the origin and dynamics of the polarization singularities in nonlinear optical processes are virtually absent. The present chapter focuses on the study of formation of *L*-type *C*-type polarization singularities in the signal beam cross section generated in various nonlinear optical processes in media with a spatial dispersion.

Many results in polarization optics were obtained within the framework of slowly varying envelope approximation, which is not suitable for the description of

the propagation of ultrashort pulses. In the description of ultrashort pulses the Stokes parameters, as well, as the orientation and the ellipticity degree of the polarization ellipse well known in “conventional” nonlinear polarization optics have no physical meaning. In this case, one should directly consider the hodograph of the electric field vector. However, it is difficult to discuss any changes of state of polarization along the signal pulse.

The conditions of appearance and the behavior of polarization singularities in the cross-section of light beam arising due to nonlinear interaction of elliptically polarized laser beams with a medium with nonlocality of quadratic and cubic optical responses are discussed in this chapter. The formation dynamics and propagation features of C -points, including pairwise creation and annihilation, for sum-frequency and second harmonic generation, beams self-action and interaction and other nonlinear optical processes are presented. The effects accompanying the propagation of ultrashort (several oscillations) elliptically polarized light pulses in nonlinear isotropic gyrotropic medium with frequency dispersion are also discussed.

9.2 Elliptically Polarized Laser Beam and Polarization Singularities

9.2.1 Formal Characterization of Elliptically Polarized Laser Beam

The well-known concepts of the “transverse” and “longitudinal” components of the electric field of monochromatic electromagnetic wave are convenient only in the case of plane waves. The beam can be naturally represented as a superposition of the plane wave spatial Fourier harmonics whose wave vectors are slightly non-collinear, and therefore even in the case when the field in the beam is directed perpendicular to its axis, it should have nevertheless a small longitudinal component. The natural generalization of these two concepts is the beams of the so-called vortex and potential types. They are specified by the conditions $\text{div}\mathbf{E}=0$ and $\text{rot}\mathbf{E}=0$, respectively. One can easily see that in this case of such a definition, each spatial Fourier harmonic of the potential or vortex beam is the longitudinal or transverse plane wave, respectively, and, therefore, all the specific properties of longitudinal or transverse electromagnetic waves will be “inherited” by the potential or vortex beams, respectively.

Note that the potential beam has a small field component directed perpendicular to its axis, while the vortex beam has a small longitudinal component. These components are the quantities of the first order of smallness in the divergence angle of the beam. It is known [15] that the potential beam cannot propagate freely and exists only inside a medium, being “coupled” with the corresponding polarization wave of matter. Having approached the surface, it makes, due to boundary

conditions, the contribution to a free wave; however, this contribution should be taken into account as a part of the signal from the surface.

The perpendicular to beam axis z part of the vortex component of electric field is completely characterized by two complex or four real parameters. It is possible select slowly varying amplitudes of circularly polarized field components $E_{\pm} = E_x \pm iE_y$ or Stokes parameters $S_0 = (|E_+|^2 + |E_-|^2)/2$, $S_1 = \text{Re}\{E_+ E_-^*\}$, $S_2 = \text{Im}\{E_+ E_-^*\}$, and $S_3 = (|E_+|^2 - |E_-|^2)/2$, where $E_{x,y}(x, y, z)$ are the slowly varying amplitudes of two orthogonal transverse complex components of the electric field. In this chapter it is convenient to characterize the incident light beam by the normalized intensity $I(x, y, z) = (|E_+|^2 + |E_-|^2)/2$, which also depends on the transverse coordinates x and y , the ellipticity degree of the polarization ellipse, $M(x, y, z) = (|E_+|^2 - |E_-|^2)/(|E_+|^2 + |E_-|^2)$, the inclination angle of its major axis, $\Psi(x, y, z) = 0.5\text{Arg}\{E_+ E_-^*\}$, and the angle $\Phi(x, y, z) = \text{Arg}\{E_+ + E_-^*\} - 0.5\text{Arg}\{E_+ E_-^*\}$, which determine the orientation of the electric field vector at the instant t measured, for example, from the principal axis of the polarization ellipse. The end of the electric field vector moves not uniformly over the polarization ellipse. The rotation direction of the electric field vector is determined by the sign of M . In homogeneously polarized beams, M , Ψ and Φ are independent of transverse coordinates. Recall that M changes from -1 (left-hand circular polarization) to 1 (right-hand circular polarization) by passing through zero (linear polarization), while Ψ from 0 to π (the states 0 and π are equivalent).

Different distributions of the light field polarization in the plane perpendicular to the beam propagation axis can be most conveniently and clearly illustrated by polarization ellipses constructed at different points of the beam cross section. The sum of squares of the semiaxes of the polarization ellipse is proportional to the light intensity at its center, the axial ratio is uniquely determined by the parameter M , and the tilt angle of its principal axis is equal to the angle Ψ . The point at the edge of each of them specifies the orientation of the electric field strength vector at the fixed instant of time. The open and filled ellipses indicate that the electric field at ellipse center is rotated clockwise and counterclockwise, respectively. Figure 9.1 shows the examples of polarization distributions in the cross section of homogeneously (a) and inhomogeneously (b) polarized beams.

For example, the electric field strength of homogeneously elliptically polarized vortex Gaussian beam propagating coincide with the z axis in vacuum is given by the formula

$$\mathbf{E}(x, y, z, t) = \left[\mathbf{e} + \frac{i\mathbf{e}_z(\mathbf{e} \cdot \nabla_{\perp})}{k} \right] \frac{E_0}{\beta(z)} \exp\left(-\frac{x^2 + y^2}{w^2 \beta(z)} - i\omega t + ik(z - l_0) \right). \quad (9.1)$$

Here, \mathbf{e}_z is the unit vector parallel to the z axis, $\nabla_{\perp} = \{\partial/\partial x, \partial/\partial y\}$, E_0 is the amplitude; ω is the frequency, w is the incident-beam half-width, the $k = \omega/c = 2\pi/\lambda$ is the wave-vector magnitude; $\beta(z) = 1 + i(z - l_0)/L_d^2$, $L_d = kw^2/2$ is the diffraction length, λ is the wavelength, c is the velocity of light. The plane of beam waist given by the expression (9.1) coincides with the plane $z = l_0$.

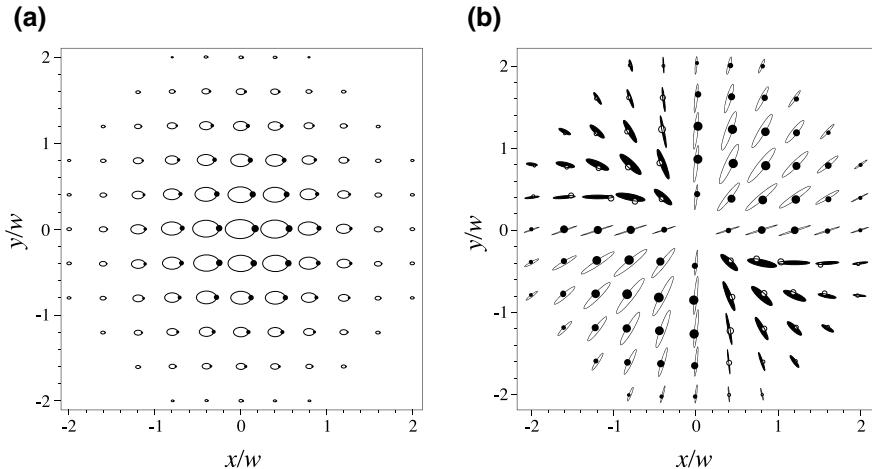


Fig. 9.1 Polarization distributions in the cross section of homogeneously (a) and inhomogeneously (b) polarized beams

Expression (9.1) contains the longitudinal component of electric field strength in the form for which in the first approximation in the parameter λ/w (the angle of divergence of beam is assumed small, $w \gg \lambda$) the necessary condition $\text{div}\mathbf{E} = 0$ is fulfilled. The unit vector \mathbf{e} without the loss of generality can be written in the form

$$\mathbf{e} = \left[(1 - M_0)^{1/2} \exp(-i\Psi_0) \mathbf{e}_+ + (1 + M_0)^{1/2} \exp(i\Psi_0) \mathbf{e}_- \right] / \sqrt{2}, \quad (9.2)$$

the $\mathbf{e}_{\pm} = \mathbf{e}_x \pm i\mathbf{e}_y$ are complex unit vectors characterizing the state of circular polarization, $|\mathbf{e}_+|^2 = |\mathbf{e}_-|^2 = 1$, $(\mathbf{e}_+ \mathbf{e}_-^*) = 0$, $\mathbf{e}_{x,y}$ are the unit vectors directed along the x and y axes. The parameter M_0 is the degree of ellipticity of radiation in homogeneously polarized beam. The angle Ψ_0 is given by the orientation of the principal axis of polarization ellipse. The polarization distributions in the cross section of this beam look like shown in Fig. 9.1.

9.2.2 Brief Information About Polarization Singularities in Laser Beams

Polarization singularities of the electric field of the light beam (points or lines in the cross section of a propagating beam in which the intensity of one of its orthogonally polarized components of the electric field becomes zero) are the subject of numerous theoretical and experimental studies in linear optics. In one of the first papers [16], the terminology extensively used at present was defined, according to which the locus of points in space where the propagating wave is circularly

(linearly) polarized, became known as *C*-lines (*L*-surfaces). The latter in the beam cross section become *C*-points and *L*-lines. Different from optical vortices, or spiral phase dislocations (where the intensity of the “scalar” field is zero) investigated usually in the approximation of constancy of polarization of propagating wave, the *C*-points, where the orientation of the polarization ellipse of the electric field strength of the radiation is not defined, can be called ‘component’ optical vortices. In their vicinity, there can be three types of morphological distributions of the polarization ellipses of the light field, called in the literature as “*star*”, “*lemon*”, and “*monster*” (see [17]).

Polarization singularity is characterized by its topological index (topological charge). It is equal to the total winding number of the polarization ellipse during one full counter-clockwise loop around the point of singularity. The charge is considered to be negative, if the rotation of the ellipse is clockwise. *C*-points of the least possible absolute value of topological charge ($1/2$) are stable under slight perturbations of the electric field, while polarization singularities with greater charges ($n/2$) split under the small perturbations of the field into $|n|$ *C*-points, each having the topological charge equal to $(1/2)\text{sgn}(n)$. While analyzing different cross sections of the light beam, which correspond to increasing values of the propagation coordinate, it is possible to observe continuous motion of *C*-points. In places where the *C*-line is tangential to the cross-section plane of propagating beam the pairwise creation or annihilation of *C*-points with opposite topological charges takes place.

As an example the trajectories of the *C*-points in a linear medium in the case of equal signs of the topological indices of the incident complex beam and the transversal polarization distribution in the propagating light are shown in Fig. 9.2. Circles indicate the *C*-line in Fig. 9.2a (*C*-points in Fig. 9.2b) with topological index equal to $1/2$, and squares indicate the *C*-line in Fig. 9.2a (*C*-points in Fig. 9.2b) with topological index equal to $-1/2$. A star designates the point of pairwise creation (annihilation) of the *C*-points. In this figure after traversing a distance the central *C*-point “exchanges” its topological charge with one of the other two *C*-points. In other words, two *C*-points with opposite topological charges annihilate, and afterwards another pair of the “oppositely charged” *C*-points appears immediately.

The conditions of appearance and the behavioral features of the light polarization singularities have been investigated in different problems of linear optics. As examples we will mention the evolution of the random vector fields in isotropic media [18], propagation of laser light in birefringent chiral crystals [19], polarization singularities in sky in daylight, propagation of optical vortices in birefringent crystals [20, 21], and coherent interaction of orthogonally polarized Bessel beams [22]. The authors of [23] established the fundamental correlation of the anisotropy of the Stokes parameters in the vicinity of the *C*-points with the morphology of the last ones. Different mechanisms leading to the appearance of the different types of polarization singularities in beams with initially regular intensity and polarization distribution propagating in inhomogeneous linear media, waveguides and laser resonators have been intensively investigated. High-performance experimental methods for the detection of polarization and phase singularities in

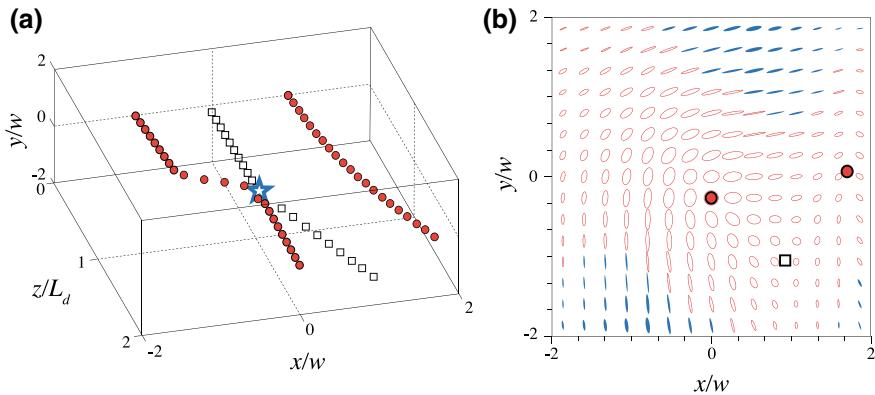


Fig. 9.2 (Adopted from [80]) The trajectories of the C-points in linear medium in the case of three singularities in incident beam (a) and the transversal polarization distribution in the propagating light at (b). Circles indicate the C-line in (a) (C-points in (b)) with 1/2-topological index, and squares indicate the C-line in (a) (C-points in (b)) with -1/2 topological index. A star designates the point of pairwise creation (annihilation) of the C-points

light beams have been developed [20, 24–26]. A review of these and other relevant works appears in [27]. In spite of the thorough studies of polarization singularities in a broad range of problems of linear optics, there exist only a very limited number of articles concerning the appearance and the evolution of polarization singularities in nonlinear optical processes. The features of creation or annihilation of polarization singularities in different nonlinear media and its evolution in remained unexplored for a long time because the account of a polarization state's changes in a propagating wave leads to rather cumbersome analytical and numerical calculations.

9.2.3 Hodograph of the Electric Field Strength of Elliptically Polarized Ultrashort Laser Pulse

Different modifications of the slowly varying envelope approximation are widely used for the description of the propagation of long pulse. In this case for the fixed time value, the spatial distribution (or temporal distribution for the fixed propagation coordinate value) of the electric field in a pulse is attributed to the set of a relatively big amount of the polarization ellipses characterizing the polarization state of radiation in different points in space (or at different moments in time). The ellipticity degree $M(t, x, y, z)$ of polarization ellipse and the angle of its orientation $\Psi(t, x, y, z)$ for each of these ellipses at certain z (or t) can be found by the same formulas as $M(x, y, z)$ and $\Psi(x, y, z)$ with $E_{x,y}(t, x, y, z)$ instead of $E_{x,y}(x, y, z)$ (see Sect. 9.2.1).

When proceeding to the shorter pulses (including those we call *ultrashort*, thus indicating several-oscillation pulses) defined as above the ellipticity degree and the angle of orientation of the polarization ellipse have no physical meaning. In such a case this also concerns any other sets of four parameters, describing the intensity and the state of the polarization of the propagating radiation. The reason is that we cannot average along the set of almost identical ellipses as earlier. In the domain of ultrashort pulses it would be correct to speak in terms of the changes of modulus of the electric field strength vector $\tilde{E}(t, x, y, z) = \left(|\tilde{E}_x|^2 + |\tilde{E}_y|^2 \right)^{1/2}$ and the angle $\Psi(t, x, y, z) = \text{arctg}(\tilde{E}_y/\tilde{E}_x)^{1/2}$ it composes with the x axis of the coordinate system, where $\tilde{E}_{x,y} = \text{Re}\{E_{x,y}(t, x, y, z)\}$. These two values wholly characterize the propagating laser pulse. Additionally, one can define quantities $M_1(t, x, y, z)$ and $\Psi_1(t, x, y, z)$, which, at certain degree, can be treated as the ellipticity degree of polarization ellipse and the angle of its orientation defined similar for the slowly varying envelope and will carry the information on the predominant orientation of the electric field strength vector. Naturally this can be done in different ways. It is only necessary to provide that these newly defined quantities $M_1(t, x, y, z)$ and $\Psi_1(t, x, y, z)$ would tend to $M(t, x, y, z)$ and $\Psi(t, x, y, z)$ correspondingly when increasing the pulse duration, and turned into $M(x, y, z)$ and $\Psi(x, y, z)$ in the extreme case of the plane monochromatic wave. The need for definition of M_1 and Ψ_1 is justified solely by the necessity of the comparison of results obtained numerically with intensity and polarization data obtained within the framework of slowly varying envelope approximation for long pulses.

In the case of ultrashort pulse it is more convenient to represent the changes of the $E_{x,y}(t, x, y, z)$ by the hodograph of the electric field vector instead of M_1 and Ψ_1 [28], because in some fragments of the propagating pulse the changes of its Cartesian components cannot be treated as similar (even roughly) to the harmonic oscillations. The hodograph can be represented as a curve in the space of parameters \tilde{E}_x , \tilde{E}_y , and z , traced by the end of the electric field vector. As an example, the hodographs of the electric field vector of the pulse output from an optically active medium are shown in Fig. 9.3 for the case when the incident pulse with Gaussian envelope has duration of several oscillation periods. After a certain distance the pulse becomes asymmetric (Fig. 9.3a), high frequency components of the pulse spectrum pass ahead of low frequency components, making the effect of the non-locality much stronger at the front of the pulse rather than in its tail. The hodograph looks like severely deformed helix with changing radius and it is not possible to discuss the polarization state of such light.

Sometimes the helicity of the hodograph changes along the pulse (for example, from right to left, or vice versa), which reflects the change of the sense of rotation of the electric field vector. Such a case is well illustrated in Fig. 9.3b, where the fragmentation of the pulse into two parts occurred, and the senses of rotation of the electric field vector within these parts are opposite.

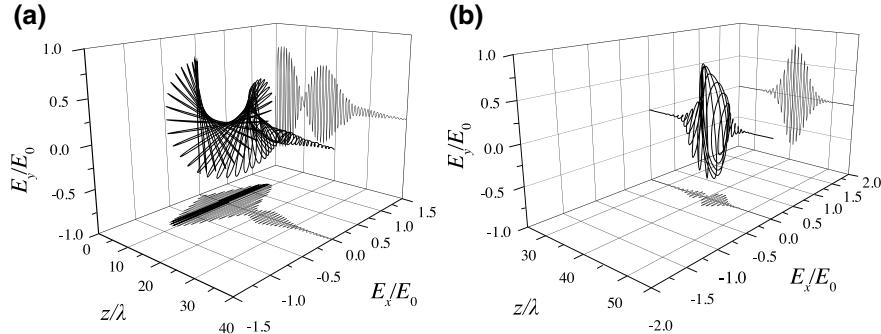


Fig. 9.3 (Adopted from [28] (a) and [82] (b)) The examples of hodograph of the electric field vector in media with frequency and spatial dispersion for short (a) and ultrashort (b) pulses. The polarizations of incident Gaussian pulses are close to linear

9.3 Nonlinear Electrodynamics of Homogeneous Media with Spatial Dispersion

9.3.1 Maxwell's Equations in Nonlinear Media with Spatial Dispersion

For the analysis of the interaction of radiation with matter, we (in common with all authors) start from the Maxwell system of equations for electromagnetic field in the medium, which directly follows from the traditional averaging of the microscopic equations:

$$\text{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad \text{rot} \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c} \mathbf{j}, \quad \text{div} \mathbf{E} = 4\pi\rho, \quad \text{div} \mathbf{B} = 0. \quad (9.3)$$

Here \mathbf{j} is the density of current induced in the medium (polarization current), ρ is the density of bound charge and $\partial\rho/\partial t + \text{div} \mathbf{j} = 0$. We also assume no charge density or density of current from external sources. The physical sense of the electric field strength \mathbf{E} and of magnetic induction \mathbf{B} entering into (9.3) is uniquely determined by the expression for Lorentz force acting on a moving point charge. Obviously the system (9.3) is not a closed one, as one assume the ρ and \mathbf{j} to be dependent on \mathbf{E} and \mathbf{B} , but the form of this dependence is not used and is not specified. The current \mathbf{j} is often represented by a sum of two components:

$$\mathbf{j} = \partial \tilde{\mathbf{P}} / \partial t + c \text{rot} \mathbf{M}, \quad (9.4)$$

and introduces vectors of electric field induction $\tilde{\mathbf{D}}$ and magnetic field strength \mathbf{H} :

$$\tilde{\mathbf{D}} = \mathbf{E} + 4\pi\tilde{\mathbf{P}}, \quad \mathbf{H} = \mathbf{B} - 4\pi\mathbf{M}. \quad (9.5)$$

In that case, the Maxwell equations (9.3) take a symmetric form

$$\text{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}, \quad \text{rot}\mathbf{H} = \frac{1}{c}\frac{\partial\tilde{\mathbf{D}}}{\partial t}, \quad \text{div}\tilde{\mathbf{D}} = 0, \quad \text{div}\mathbf{B} = 0, \quad (9.6)$$

and the material equations $\tilde{\mathbf{D}} = \tilde{\mathbf{D}}(\mathbf{E}, \mathbf{B})$ and $\mathbf{H} = \mathbf{H}(\mathbf{E}, \mathbf{B})$ can be written down also in symmetric form. The approach depicted above is called “symmetric” [29]. It has a number of significant shortcoming. The relations (9.4) and (9.5) are not unique since they introduce four new quantities ($\tilde{\mathbf{P}}$, \mathbf{M} , $\tilde{\mathbf{D}}$ and \mathbf{H}) using only three relationships. Sometimes by writing (9.4) one says that $\tilde{\mathbf{P}}$ and \mathbf{M} are connected with the electric and magnetic moments of the medium, respectively. This, however, is not quite correct, as it remains unclear to that degree these quantitative definitions of $\tilde{\mathbf{P}}$ and \mathbf{M} determine them qualitatively, because in the optical frequency range the notion of the magnetic moment of the medium with nonlocality of optical response (spatial dispersion) loses its physical meaning. At the phenomenological level of consideration, it is quite artificial to split \mathbf{j} into two terms in an arbitrary way, then to write down the constitutive relation for each term, and finally to find the dependence of the total polarization current on the electric field strength. This last dependence is very important because the total polarization current enter into the right side of the wave equation for \mathbf{E} . Knowing electric field strength it is easy to find \mathbf{B} . The complexity of consistent generalization to the case of nonlinear media is the other significant defect of the “symmetric approach” to electrodynamics.

It is most natural to proceed from system (9.3) directly and to write down the material equation for the total polarization current \mathbf{j} as $\mathbf{j} = \partial\mathbf{P}/\partial t$, where \mathbf{P} is the total medium polarization (Landau–Lifshitz approach) [30, 31]. In this case the corresponding electric field induction $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ and magnetic field strength $\mathbf{H} = \mathbf{B}$. It is practically the same to consider $\mathbf{M} = 0$ in “symmetric approach”. Note that $\mathbf{P} \neq \tilde{\mathbf{P}}$ and $\mathbf{D} \neq \tilde{\mathbf{D}}$. Vector $\mathbf{j}(\mathbf{r}_0, t)$ and hence $\mathbf{P}(\mathbf{r}_0, t)$ and $\mathbf{D}(\mathbf{r}_0, t)$ can, in general, depend on the fields \mathbf{E} and \mathbf{B} not only in the point \mathbf{r}_0 , but at the adjacent points as well. On the other hand, the field $\mathbf{B}(\mathbf{r}, t)$ is related to the field $\mathbf{E}(\mathbf{r}, t)$ by the first equation of the system (9.3). This allows one to consider, in phenomenological treatment, that $\mathbf{P}(\mathbf{r}_0, t)$ and $\mathbf{D}(\mathbf{r}_0, t)$ depend only on the field $\mathbf{E}(\mathbf{r}, t)$ in the entire space. In this approach to electrodynamics of media with spatial dispersion, the Maxwell system of equations and the material equations $\mathbf{D} = \mathbf{D}(\mathbf{E}, \mathbf{B})$ and $\mathbf{H} = \mathbf{H}(\mathbf{E}, \mathbf{B})$ take the form

$$\text{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}, \quad \text{rot}\mathbf{B} = \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}, \quad \text{div}\mathbf{D} = 0, \quad \text{div}\mathbf{B} = 0, \quad (9.7)$$

$$\mathbf{D}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) + 4\pi\mathbf{P}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, t) + 4\pi \sum_{n=1}^{\infty} \mathbf{P}^{(n)}(\mathbf{r}, t), \quad (9.8)$$

where

$$\begin{aligned} \mathbf{P}^{(n)}(\mathbf{r}, t) = & \int_{-\infty}^t dt_1 \int_{-\infty}^t dt_2 \dots \int_{-\infty}^t dt_n \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \\ & \dots \int d\mathbf{r}_n \hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n); \mathbf{E}(\mathbf{r}_1, t_1) \mathbf{E}(\mathbf{r}_2, t_2) \dots \mathbf{E}(\mathbf{r}_n, t_n). \end{aligned} \quad (9.9)$$

It is necessary to stress that by assuming $\mathbf{B} = \mathbf{H}$ one does not neglect by any magnetic effects at all. All are taken into account by the relation (9.9), which, in particular, can be reflected in symmetry properties of the tensors $\hat{\chi}^{(n)}$ of the $n + 1$ rank, which contain all the information about the linear and nonlinear (including magnetic) optical properties of the matter. Formally, in (9.9) the integration is taken over the entire range x_i, y_i, z_i , where $\mathbf{r}_i = \{x_i, y_i, z_i\}$, $i = 1, 2, \dots, n$. However, in matter with weak spatial dispersion the material tensors $\hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ are noticeably distinct from zero only in the small region of space about the point determined by vector \mathbf{r} with linear dimensions $d \ll \lambda$. Within the bounds of discussed approach, the boundary conditions and material equations for electric and magnetic fields cannot be considered independently of each other. In accordance with the logic of electrodynamics first come the Maxwell equation, then the material equations, and only after those boundary conditions, expressions for energy, etc., which are consequences of the Maxwell equations and the material relations and thus vary the latter.

9.3.2 Material Equation in Nonlinear Media with Spatial Dispersion. Local and Nonlocal Susceptibilities

For stationary homogeneous media material tensors $\hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ depend only on $t - t_n$ and $\mathbf{r} - \mathbf{r}_n$ and the expression (9.9) can be written for components $\mathbf{E}_{i_m}(\omega_i, \mathbf{r}_i)$ and $\mathbf{P}_i^{(n)}(\omega, \mathbf{r})$ of temporal Fourier transforms of the electric field $\mathbf{E}(t_i, \mathbf{r}_i)$ and the medium polarization $\mathbf{P}^{(n)}(\mathbf{r}, t)$ ($m = 1, 2, \dots, n$) in the following way:

$$\begin{aligned} P_i^{(n)}(\omega, \mathbf{r}) = & \int_{-\infty}^{\infty} \delta(\omega - \omega_1 - \omega_2 - \dots - \omega_n) d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \dots \int_{-\infty}^{\infty} d\omega_n \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \\ & \int d\mathbf{r}_n \tilde{\chi}_{i_1 i_2 \dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2, \dots, \mathbf{r} - \mathbf{r}_n) E_{i_1}(\omega_1, \mathbf{r}_1) E_{i_2}(\omega_2, \mathbf{r}_2) \dots E_{i_n}(\omega_n, \mathbf{r}_n), \end{aligned} \quad (9.10)$$

where the indices i and i_n number the coordinates of the vectors and take the values x, y and z . Here and below the summation over repeated indices is performed, δ is

Dirac delta function. In (9.10) $\tilde{\chi}_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2, \dots, \mathbf{r} - \mathbf{r}_n)$ is the temporal Fourier transforms of $\hat{\chi}^{(n)}(t - t_1, t - t_2, \dots, t - t_n, \mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2, \dots, \mathbf{r} - \mathbf{r}_n)$. Additional frequency argument $\omega = \omega_1 + \omega_2 + \dots + \omega_n$ is assigned formally to the first index i of the tensor $\tilde{\chi}_{ii_1i_2\dots i_n}^{(n)}$.

Electric field $\mathbf{E}_{i_m}(\omega_m, \mathbf{r}_m)$ can be represented in the form of Taylor expansion at the point determined by vector \mathbf{r} , in which one holds only the first two members due to weak spatial dispersion of the medium under study ($|\mathbf{r} - \mathbf{r}_m| \ll r$)[30, 31]. After substituting

$$\mathbf{E}_{i_m}(\omega_m, \mathbf{r}_m) \approx \mathbf{E}_{i_m}(\omega_m, \mathbf{r}) + (\mathbf{r}_m - \mathbf{r}) \frac{\partial \mathbf{E}_{i_m}(\omega_m, \mathbf{r})}{\partial \mathbf{r}} \quad (9.11)$$

in (9.10) the material equation $\mathbf{P}(\mathbf{E})$ can be written in the following form:

$$\begin{aligned} P_i^{(n)}(\omega, \mathbf{r}) = & \frac{1}{(2\pi)^{(n-1)}} \int_{-\infty}^{\infty} \delta(\omega - \omega_1 - \omega_2 - \dots - \omega_n) d\omega_1 \int_{-\infty}^{\infty} d\omega_2 \dots \int_{-\infty}^{\infty} d\omega_n \\ & \times \left[\chi_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r}) E_{i_1}(\omega_1, \mathbf{r}) E_{i_2}(\omega_2, \mathbf{r}) \dots E_{i_n}(\omega_n, \mathbf{r}) \right. \\ & + \sum_{m=1}^n \gamma_{ii_1i_2\dots i_n l}^{(n)}(\omega; \omega_m, \omega_2 - \delta_{2m}, \omega_3 - 2\delta_{3m}, \omega_4 - 3\delta_{4m}, \dots, \omega_{n-(n-1)\delta_{nm}}, \mathbf{r}) (\nabla_l E_{i_1}(\omega_m, \mathbf{r})) \\ & \left. \times E_{i_2}(\omega_2 - \delta_{2m}, \mathbf{r}) E_{i_3}(\omega_3 - 2\delta_{3m}, \mathbf{r}) E_{i_4}(\omega_4 - 3\delta_{4m}, \mathbf{r}) \dots E_{i_n}(\omega_{n-(n-1)\delta_{nm}}, \mathbf{r}) \right]. \end{aligned} \quad (9.12)$$

Here $l = x, y, z$, $\nabla_x = \partial/\partial x$, $\nabla_y = \partial/\partial y$, $\nabla_z = \partial/\partial z$. Let's call the tensors of the $n+1$ rank

$$\begin{aligned} \chi_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r}) = & \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \\ \dots \int d\mathbf{r}_n \tilde{\chi}_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2, \dots, \mathbf{r} - \mathbf{r}_n) \end{aligned} \quad (9.13)$$

“local optical susceptibilities”, and the tensors of the $n+2$ rank

$$\begin{aligned} \gamma_{ii_1i_2\dots i_n l}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r}) = & \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \\ \dots \int d\mathbf{r}_n \tilde{\chi}_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r} - \mathbf{r}_1, \mathbf{r} - \mathbf{r}_2, \dots, \mathbf{r} - \mathbf{r}_n) d'_l \end{aligned} \quad (9.14)$$

“nonlocal optical susceptibilities”. In (9.14) $d'_x = x_1 - x$, $d'_y = y_1 - y$, and $d'_z = z_1 - z$. Local susceptibilities $\chi_{ii_1i_2\dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r})$ are symmetric with respect to the permutation of all indices except the first one with simultaneous rearrangement of the frequency arguments. Nonlocal susceptibilities are symmetric with respect to the permutation of all indices except the first, the second and the last indices with

simultaneous rearrangement of the frequency arguments. Tensors (9.14) are a consequence of the presence of the nonlocality of nonlinear optical response of the medium under study. The relative contribution of nonlocal members in the expression for generalized polarization is usually small as compared with local contribution of the same order. For this reason, they can often be neglected. However, in situations when the local susceptibilities vanish due to medium symmetry or frequency equality, the nonlocal susceptibilities start to play important role. It is a nonlocal susceptibilities determine in this case the intensity and the polarization state of the signal waves arising in the process of the interaction of incident radiation with a nonlinear medium.

If n monochromatic waves collinearly propagate in the medium then, taking into account the symmetry properties of $\chi_{ii_1 i_2 \dots i_n}^{(n)}$ and $\gamma_{ii_1 i_2 \dots i_n l}^{(n)}$, we can write down (9.12) in the following form:

$$\begin{aligned} P_i^{(n)}(\omega, \mathbf{r}) = & D_n \chi_{ii_1 i_2 \dots i_n}^{(n)}(\omega; \omega_1, \omega_2, \dots, \omega_n, \mathbf{r}) E_{i_1}(\omega_1, \mathbf{r}) E_{i_2}(\omega_2, \mathbf{r}) \dots E_{i_n}(\omega_n, \mathbf{r}) \\ & + \sum_{m=1}^n D_{n-1}(\omega_m) \gamma_{ii_1 i_2 \dots i_n l}^{(n)}(\omega; \omega_m, \omega_{2-\delta_{2m}}, \omega_{3-2\delta_{3m}} \dots, \omega_{n-(n-1)\delta_{nm}}, \mathbf{r}) (\nabla_l E_{i_1}(\omega_m, \mathbf{r})) \\ & \times E_{i_2}(\omega_{2-\delta_{2m}}, \mathbf{r}) E_{i_3}(\omega_{3-2\delta_{3m}}, \mathbf{r}) \dots E_{i_n}(\omega_{n-(n-1)\delta_{nm}}, \mathbf{r}). \end{aligned} \quad (9.15)$$

Here $D_n = D(\omega_1, \omega_2, \dots, \omega_n) = n!$ if all n frequencies are different, $D_n = (n-1)!$ if only two frequencies from $\omega_1, \omega_2, \dots, \omega_n$ coincide, $D_n = (n-2)!$ if only three frequency are equal to each other, $D_{n-1}(\omega_m) = D(\omega_1, \omega_2, \dots, \omega_{m-1}, \omega_{m+1}, \dots, \omega_n)$, and $D_1 = 1$. Information about the structure of nonlinear optical susceptibilities (both local and nonlocal) up to tensors of seven rank for all crystal classes and limiting groups can be found in [32–34]. The examples of specific form of (9.15) for different nonlinear optical phenomena we consider below.

9.3.3 Boundary Conditions for Electromagnetic Field and Surface Nonlinear Optical Response. Surface Susceptibilities

The main problem that arises in deriving boundary conditions at the intermedium surface in the case of media with spatial dispersion is in taking into account correctly the influence of thin transition layers on light reflection. Such layers may arise on the surface of any media; however, their influence is stronger on the surface of media possessing spatial dispersion. The reason is that their characteristic spatial scale d_0 cannot be less than the scale d of nonlocal optical response of the medium. Conventional boundary conditions are derived from Maxwell equations by passing to the limit, i.e., under the assumption $d_0 = 0$ (the sharp boundary between uniform media).

This contradiction can be resolved by making the assumption [35, 36] that in a transition layer all medium characteristics and, hence, of the electromagnetic field vary continuously although rather rapidly. By solving approximately the Maxwell equations in the transition layer in the required approximation with respect to a small parameter $d_0/\lambda \approx d/\lambda$ it is possible to find relationships between the normal and tangential components of the electric field strength, the magnetic induction, the electric field induction, and the magnetic field strength on its opposite boundaries. By comparing the relationships obtained in this way with those corresponding to the case of a sharp boundary between uniform media one can write down corrected boundary conditions capable, however, of taking into account transitional layer contribution on light reflection and refraction. All additional terms in the corrected boundary conditions can be considered as due to the existence of an additional surface polarization current \mathbf{i} on the sharp boundary of the two media (existing only on the intermedium surface). This procedure was realized in [35–37].

In the first approximation in d_0/λ these conditions relating the vector components of the strength and inductions of the electric and magnetic field at the frequency ω in vacuum (the superscript “ v ”) and in matter with the nonlocality of the nonlinear optical response (the superscript “ t ”) have at the plane interface the following form:

$$\begin{aligned} E_{\tan}^{(v)}(\omega) - E_{\tan}^{(t)}(\omega) &= (-4\pi i/\omega) \text{grad}_{\tan}[i_n(\omega)], \\ D_n^{(v)}(\omega) - D_n^{(t)}(\omega) &= (-4\pi i/\omega) \text{div}[\mathbf{i}_{\tan}(\omega)], \\ \mathbf{B}^{(v)}(\omega) - \mathbf{B}^{(t)}(\omega) &= (-4\pi/c)[\mathbf{n} \times \mathbf{i}(\omega)]. \end{aligned} \quad (9.16)$$

Here \mathbf{n} is the surface-perpendicular unit vector directed opposite to the z axis from the medium to the vacuum, subscripts “ \tan ” and “ n ” denote tangential and normal components of the vectors \mathbf{E} , \mathbf{D} , \mathbf{B} , and \mathbf{i} . The latter can be interpreted as a surface current density of coupled charges [36, 37] (surface polarization current) and in the general case in the first approximation in the parameter d_0/λ it can be represented in the framework of the phenomenological approach in the form of expansion in powers $\mathbf{E}^{(v)}$:

$$\mathbf{i} = \hat{\kappa}^{(1)} \cdot \mathbf{E}^{(v)} + \hat{\kappa}^{(2)} : \mathbf{E}^{(v)} \mathbf{E}^{(v)} + \hat{\kappa}^{(3)} : \mathbf{E}^{(v)} \mathbf{E}^{(v)} \mathbf{E}^{(v)} + \dots, \quad (9.17)$$

where the material tensors $\hat{\kappa}^{(n)}$ characterize the surface response of the nonlinear medium to external electromagnetic field $\mathbf{E}^{(v)}$. The tensors $\hat{\kappa}^{(n)}$ are material tensors in the same sense as $\hat{\chi}^{(n)}$ and $\hat{\gamma}^{(n)}$, i.e., they are independent of the parameters of the incident radiation. From formulas (9.17) it follows that in a first approximation in the spatial dispersion parameter all the features of light reflection and refraction at the plane surface of nonlinear media with nonlocal optical response and near-surface inhomogeneity are determined by the surface polarization current \mathbf{i} . It should be introduced into consideration if we use the model of a sharp intermedium boundary between vacuum and nonlinear medium. It is necessary to stress that all

the components of the tensor $\hat{\kappa}^{(n)}$ are definitely connected with the tensor $\hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$. When we pass from tensors $\hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ to the tensors $\hat{\chi}^{(n)}(t-t_1, t-t_2, \dots, t-t_n, \mathbf{r}-\mathbf{r}_1, \mathbf{r}-\mathbf{r}_2, \dots, \mathbf{r}-\mathbf{r}_n)$, characterizing the thickness of homogeneous medium, part of the information contained in the full tensor $\hat{\chi}^{(n)}(t, t_1, t_2, \dots, t_n, \mathbf{r}, \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ is lost. Tensor $\hat{\kappa}^{(n)}$ takes up it.

9.3.4 Basis List of Nonlocal Nonlinear Optical Effects

Local, nonlocal and surface nonlinear susceptibilities determine all known nonlinear optical phenomena. For example, using general expression (9.14) one can easily find that the quadratic to the electric field part of nonlinear medium polarization at frequency $\omega_1 \pm \omega_2$ consists of three items:

$$\begin{aligned} P_i^{(2)}(\omega_1 \pm \omega_2, \mathbf{r}) = & 2\chi_{ii_1 i_2}^{(2)}(\omega_1 \pm \omega_2; \omega_1, \pm \omega_2, \mathbf{r}) E_{i_1}(\omega_1, \mathbf{r}) E_{i_2}(\pm \omega_2, \mathbf{r}) \\ & + \gamma_{ii_1 i_2 l}^{(2)}(\omega_1 \pm \omega_2; \omega_1, \pm \omega_2, \mathbf{r}) (\nabla_l E_{i_1}(\omega_1, \mathbf{r})) E_{i_2}(\pm \omega_2, \mathbf{r}) \\ & + \gamma_{ii_1 i_2 l}^{(2)}(\omega_1 \pm \omega_2; \pm \omega_2, \omega_1, \mathbf{r}) (\nabla_l E_{i_2}(\pm \omega_2, \mathbf{r})) E_{i_1}(\omega_1, \mathbf{r}), \end{aligned} \quad (9.18)$$

which describe the specific features of sum and difference frequency generation processes in the medium bulk. In addition, surface susceptibility $\hat{\kappa}^{(2)}(\omega_1 \pm \omega_2; \omega_1, \pm \omega_2)$ also determines the light intensity and polarization upon sum-frequency and difference-frequency generation processes from the surface of nonlinear medium.

Cubic medium polarization at frequency $\omega_1 + \omega_2 + \omega_3$:

$$\begin{aligned} P_i^{(3)}(\omega_1 + \omega_2 + \omega_3, \mathbf{r}) = & D_3 \chi_{ii_1 i_2 i_3}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_1, \omega_2, \omega_3, \mathbf{r}) E_{i_1}(\omega_1, \mathbf{r}) E_{i_2}(\omega_2, \mathbf{r}) E_{i_3}(\omega_3, \mathbf{r}) \\ & + D_2 \gamma_{ii_1 i_2 l}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_1, \omega_2, \omega_3, \mathbf{r}) (\nabla_l E_{i_1}(\omega_1, \mathbf{r})) E_{i_2}(\omega_2, \mathbf{r}) E_{i_3}(\omega_3, \mathbf{r}) \\ & + D_2 \gamma_{ii_1 i_2 l}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_2, \omega_1, \omega_3, \mathbf{r}) (\nabla_l E_{i_2}(\omega_2, \mathbf{r})) E_{i_1}(\omega_1, \mathbf{r}) E_{i_3}(\omega_3, \mathbf{r}) \\ & + D_2 \gamma_{ii_1 i_2 l}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_3, \omega_1, \omega_2, \mathbf{r}) (\nabla_l E_{i_3}(\omega_3, \mathbf{r})) E_{i_1}(\omega_1, \mathbf{r}) E_{i_2}(\omega_2, \mathbf{r}), \end{aligned} \quad (9.19)$$

includes one item with local cubic susceptibility and three items with nonlocal optical susceptibilities. This expression is used for description of light self-focusing ($\omega_1 = \omega_2 = -\omega_3 = \omega$), nonlinear optical activity ($\omega_1 = \omega_2 = -\omega_3 = \omega$), third frequency generation ($\omega_1 = \omega_2 = \omega_3 = \omega$), CARS ($\omega_1 = \omega_2 = \tilde{\omega}_1, \omega_3 = -\tilde{\omega}_2$), etc.

Theoretical investigation of elliptically polarized laser beam self-action and interaction during both its propagation and reflection from a medium with spatial nonlocality of nonlinear optical response gives ability to arrive at a conclusion that the all terms in expressions for nonlinear medium polarization wake up quantitative

and qualitative changes in the state of polarization of propagating radiation. In addition, the state of polarization of incident laser beams and pulses has a great influence on the nonlinear optical processes under study.

In plane wave approximation the propagation of elliptically polarized light along optical axis in crystals of cubic, hexagonal and tetragonal crystal systems always accompanied by intensity dependent rotation and deformation of the light polarization ellipse. The expression for the nonlinear medium polarization in this case has the following form:

$$\begin{aligned} P_i^{(3)}(\omega, \mathbf{r}) = & 3\chi_{ijlm}^{(3)}(\omega; \omega, \omega, -\omega, \mathbf{r})E_j(\omega, \mathbf{r})E_l(\omega, \mathbf{r})E_m^*(\omega, \mathbf{r}) \\ & + 2\gamma_{ijlmn}^{(3)}(\omega; \omega, \omega, -\omega, \mathbf{r})(\nabla_n E_j(\omega, \mathbf{r}))E_l(\omega, \mathbf{r})E_m^*(\omega, \mathbf{r}) \\ & + 2\gamma_{ijlmm}^{(3)}(\omega; \omega, \omega, -\omega, \mathbf{r})(\nabla_n E_l(\omega, \mathbf{r}))E_j(\omega, \mathbf{r})E_m^*(\omega, \mathbf{r}) \\ & + \gamma_{ijlmm}^{(3)}(\omega; \omega, \omega, -\omega, \mathbf{r})(\nabla_n E_m^*(\omega, \mathbf{r}))E_j(\omega, \mathbf{r})E_l(\omega, \mathbf{r}), \end{aligned} \quad (9.20)$$

and four different mechanisms of intensity dependent rotation and deformation of the light polarization ellipse take place during the wave propagation [38–44]. They are connected with real and imaginary parts of local and nonlocal cubic susceptibilities included in the (9.20). The first mechanism is caused by the proper nonlinear medium gyrotropy ($\text{Re}\{\hat{\gamma}^{(3)}\}$). It takes place in fifteen crystallographic classes and Curie symmetry groups (see Table 9.1). The second mechanism arises from the anisotropy of the imaginary parts of local cubic susceptibility. It fulfils oneself in nineteen crystallographic classes. The light polarization ellipse changes its own orientation during the propagation in all media. This mechanism is associated with $\text{Re}\{\hat{\chi}^{(3)}\}$. The fourth mechanism determined by the anisotropy of the imaginary parts of nonlocal cubic susceptibility ($\text{Im}\{\hat{\gamma}^{(3)}\}$) takes place in fourteen crystallographic classes. The last two mechanisms contribute to intensity dependent rotation and deformation of the light polarization ellipse only in the case of elliptical polarization of the incident wave.

Theoretical study of polarization self-action of light propagating through a nonlinear crystal permits the creation of a new type of spectroscopy [38, 39, 42–46], which provides valuable information about cubic local, nonlocal and surface nonlinearities. Fourier analysis of the dependence of polarization ellipse deformation and rotation at the output of crystals on incident wave orientation of the polarization ellipse and its degree of ellipticity is the only available method to process the obtained experimental data about $\chi_{ijlm}^{(3)}(\omega; \omega, \omega, -\omega)$, $\gamma_{ijlmn}^{(3)}(\omega; \omega, \omega, -\omega)$ and $\kappa_{ijlm}^{(3)}(\omega; \omega, \omega, -\omega)$. This scheme of spectroscopy allows one to distinguish between contributions of different mechanisms of nonlinear optical rotation and deformation of polarization ellipse irrespective of the crystal symmetry.

The study of these effects during the interaction of elliptically polarized radiation with the crystals of cubic, hexagonal and tetragonal crystal systems is shown to give one and a half or twice as much spectroscopy information on the cubic nonlinearity

Table 9.1 Different mechanisms of intensity dependent rotation and deformation of the light polarization ellipse in crystallographic classes and Curie symmetry groups

Classes	Mechanisms			
	1	2	3	4
4, 422, 432, 3, 6, 23, ∞	+	+	+	+
32, 622, ∞ 2, ∞ ∞	+		+	
$\bar{4}$, 42 m, 4mm, $\bar{4}3$ m	+	+	+	+
4/m, $\bar{3}$, $\bar{6}$, 6/m, m3m, ∞ /m, m3, 4/mmm		+	+	
$\bar{3}$ m, 6/mmm, $\bar{6}$ m2, ∞ /mm, ∞ ∞ m			+	
3m, 6mm, ∞ m			+	+

and its spatial dispersion as in the case of application of a plane-polarized wave. Moreover the joint study of light polarization self-action during both reflection and propagation of laser radiation through the mentioned above crystal is most effective. This is connected with the fact that unique spectroscopy information received in experiments “on reflection” and “on propagation” is, as shown, complementary to a great degree.

9.3.5 *Phenomenological Theory of Nonlinear Optical Response of the Bulk of Homogeneous Isotropic Media with Broken Mirror Symmetry*

This section illustrates the strong influence of the space group of symmetry of the medium on the specific form of the mentioned above material equations (9.18)–(9.20) by the example of homogeneous isotropic medium with broken mirror symmetry. The typical example of such medium, with space group of symmetry $\infty\infty$, is macroscopically noncentrosymmetric isotropic solution of chiral (left-right asymmetric) macromolecules, which are the primary elements of living organisms. On the basis of measurements of the frequency dispersion of the nonlinear even order local optical susceptibilities of such solutions (which susceptibilities are absent in centrosymmetric liquids and racemic solutions of chiral molecules) and the nonlinear odd-order nonlocal optical susceptibilities (also present exclusively in optically active liquids and solutions) it is possible to develop new schemes for coherent nonlinear optical spectroscopy and possess enhanced information about these media.

Local and nonlocal susceptibilities of medium with space group of symmetry $\infty\infty$ are expressed through a linear combination of summands of the form $e_{ijk}\delta_{lm}$ and $\delta_{ij}\delta_{kl}$ [32], where e_{ijk} is the Levi-Civita symbol, δ_{kl} is the Kronecker delta, in the following form:

$$\chi_{ijk}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2) = \chi^{(2)}(\omega_1, \omega_2) e_{ijk}, \quad (9.21)$$

$$\chi_{ijkl}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_1, \omega_2, \omega_3) = \chi_1^{(3)} \delta_{ij} \delta_{kl} + \chi_2^{(3)} \delta_{ik} \delta_{jl} + \chi_3^{(3)} \delta_{il} \delta_{jk}, \quad (9.22)$$

$$\begin{aligned} \chi_{ijklm}^{(4)}(\omega_1 + \omega_2 + \omega_3 + \omega_4; \omega_1, \omega_2, \omega_3, \omega_4) = & \chi_1^{(4)} e_{ijk} \delta_{ml} + \chi_2^{(4)} e_{ijl} \delta_{km} + \chi_3^{(4)} e_{ijm} \delta_{kl} \\ & + \chi_4^{(4)} (e_{ikl} \delta_{jm} + e_{ikm} \delta_{jl} + e_{jkl} \delta_{im} + e_{jkm} \delta_{il}) \\ & + \chi_5^{(4)} (e_{jml} \delta_{ik} + e_{iml} \delta_{jk}) + \chi_6^{(4)} e_{kml} \delta_{ij}, \end{aligned} \quad (9.23)$$

$$\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2) = \gamma_1^{(2)} \delta_{ij} \delta_{kl} + \gamma_2^{(2)} \delta_{ik} \delta_{jl} + \gamma_3^{(2)} \delta_{il} \delta_{jk}, \quad (9.24)$$

$$\begin{aligned} \gamma_{ijklm}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_1, \omega_2, \omega_3) = & \gamma_1^{(3)} e_{ijk} \delta_{ml} + \gamma_2^{(3)} e_{ijl} \delta_{km} + \gamma_3^{(3)} e_{ijm} \delta_{kl} \\ & + \gamma_4^{(3)} (e_{ikl} \delta_{jm} + e_{ikm} \delta_{jl} + e_{jkl} \delta_{im} + e_{jkm} \delta_{il}) \\ & + \gamma_5^{(3)} (e_{jml} \delta_{ik} + e_{iml} \delta_{jk}) + \gamma_6^{(3)} e_{kml} \delta_{ij} \end{aligned} \quad (9.25)$$

Here $\chi_{1,2,3}^{(3)} = \chi_{1,2,3}^{(3)}(\omega_1, \omega_2, \omega_3)$, $\chi_{1,2,3,4,5,6}^{(4)} = \chi_{1,2,3,4,5,6}^{(4)}(\omega_1, \omega_2, \omega_3, \omega_4)$, $\gamma_{1,2,3}^{(2)} = \gamma_{1,2,3}^{(2)}(\omega_1, \omega_2)$, $\gamma_{1,2,3,4,5,6}^{(3)} = \gamma_{1,2,3,4,5,6}^{(3)}(\omega_1, \omega_2, \omega_3)$.

When two pump plane waves with frequencies $\omega_{1,2}$ and electric fields $\mathbf{E}_{1,2}$ collinearly propagate through an isotropic gyrotropic medium the main contribution to sum-frequency generation process gives the local susceptibility $\chi_{ijk}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$. It might seem that medium polarization

$$\mathbf{P}^{(2)}(\omega_1 + \omega_2) = 2\chi^{(2)}(\omega_1, \omega_2) [\mathbf{E}_1 \times \mathbf{E}_2] \quad (9.26)$$

is purely longitudinal in this case (i.e. directed along the wave vectors of propagated waves). The longitudinal polarization, in its turn, cannot produce a free electromagnetic wave and sum frequency generation does not occur in noncentrosymmetric medium. However, this conclusion is true only in a very rough approximation, when both pump beams are assumed to be plane. Early it was mentioned that each freely propagating pump beam should possess a small longitudinal component (of the first order of smallness in its divergence angle). The longitudinal components of the electric fields of fundamental beams may give rise to a vortex part of the vector $\mathbf{P}^{(2)}(\omega_1 + \omega_2)$ and, as a consequence, to a signal at sum frequency from the medium bulk. A contribution of $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$ is small enough to sum frequency generation process in this case.

The appearance of a signal at the doubled frequency due to local optical susceptibility $\hat{\chi}^{(2)}$ is forbidden ($\chi^{(2)}(\omega, \omega) \equiv 0$) not only in centrosymmetric isotropic medium (space group of symmetry $\infty\infty m$) but also in noncentrosymmetric isotropic (gyrotropic) medium (space group of symmetry $\infty\infty$), since the

corresponding tensor $\chi_{ijk}^{(2)}(2\omega; \omega, \omega) \equiv 0$ because of symmetry under permutation of its last two subscripts. Nevertheless, second harmonic generation is possible because the material equation for the medium polarization $\mathbf{P}^{(2)}(2\omega)$ contains nonlocal susceptibility $\gamma_{ijkl}^{(2)}(2\omega; \omega, \omega)$. Invariant expression can be written in the form [47]

$$\mathbf{P}^{(2)}(2\omega) = \frac{1}{2}\gamma_1^{(2)}(\omega, \omega)\vec{\nabla}(\mathbf{E} \cdot \mathbf{E}) + \gamma_2^{(2)}(\omega, \omega)\mathbf{E}(\vec{\nabla} \cdot \mathbf{E}) + \gamma_3^{(2)}(\omega, \omega)(\mathbf{E} \cdot \vec{\nabla})\mathbf{E}. \quad (9.27)$$

The first term in the right hand side of (9.27) is the gradient of a scalar function, and therefore it gives a purely irrotational contribution to the medium polarization. Created by this part of $\mathbf{P}^{(2)}(2\omega)$ electric field is also irrotational type, which, as is well known, cannot propagate freely but exists only inside a medium. When this electric field encounters the surface of the medium, it contributes to a free wave because the boundary conditions for an electromagnetic field must be satisfied. This contribution is taken into account as part of the signal at the double frequency from the surface. The second term in material equation (9.27) is proportional to $\text{div}\mathbf{E}$, and therefore equals zero, which follows from Maxwell's equation $\text{div}\mathbf{D} = 0$ taking into account the constant-pump approximation. Thus, the second harmonic signal can be generated only by the solenoidal component of the third term of (9.27). However, if we consider a plane pump wave, this term equals zero because the field of a free plane wave is transverse. Unfortunately Gaussian beam (9.1) (the beam with a homogeneous distribution of the polarization state over its cross-section) also cannot generate a second harmonic signal. Its contribution to the last term in the right hand side of (9.27) equals zero.

The generation of signal wave is possible due to $\gamma_3^{(2)}(\omega, \omega)$ only if degree of ellipticity of the polarization ellipse of fundamental beam changes in the beam cross-section. The signal at the double frequency was experimentally obtained in the interior of a noncentrosymmetric solution of arabinose using noncollinear interaction of two laser beams with identical frequency ($\omega_1 = \omega_2$) [48, 49]. It was not quadratic but fourth-order in the field of the fundamental wave, which is explained by the fact that it displayed local optical susceptibility $\chi_{ijkml}^{(4)}(2\omega; \omega, \omega, \omega, -\omega) = \tilde{\chi}^{(4)}(e_{ijl}\delta_{km} + e_{ikl}\delta_{jm} + e_{iml}\delta_{jk})$, where $\tilde{\chi}^{(4)} = \chi_{2,4,5}^{(4)}(\omega, \omega, \omega, -\omega)$. This scheme of noncollinear interaction of two laser beams can be easily generalized to the nondegenerate case ($\omega_1 \neq \omega_2$), in which the five-photon process arising by the susceptibility $\chi_{ijkml}^{(4)}(3\omega - \omega_1; \omega, \omega, \omega, \omega_1)$. Such wave interaction can also be synchronous in the medium with normal frequency dispersion by a suitable choice of the angle between the wave vectors of interacting waves. In some sense this noncollinear scheme of interaction is analogous to the widely used scheme of coherent anti-Stokes Raman scattering (CARS), based on the four photon interaction ($2\omega - \omega_1 = \omega + \omega - \omega_1$). It is of a great interest for studies of chiral biomolecules in

the solution and called BioCARS [50, 51] since. The resonances in the BioCARS measurements take place if the Raman resonance condition or the hyper-Raman resonance conditions (as in ordinary CARS) are fulfilled. Let us stress that BioCARS is impossible in isotropic centrosymmetric media.

A more complete description of nonlinear optical processes in medium, with space group of symmetry $\infty\infty$ should include an account of nonlocal nonlinear susceptibilities in the general case. For example, additional contribution to the equation for $\mathbf{P}^{(3)}(\omega_1 + \omega_2 + \omega_3)$ associated with the spatial dispersion of cubic nonlinearity is very cumbersome, if $\omega_1 \neq \omega_2 \neq \omega_3$. The interesting scheme operates with the $\gamma_{ijkml}^{(3)}(\omega - \omega + \omega_1; \omega, -\omega, \omega_1)$, where ω and ω_1 are the frequencies of the pump beam and the probe beam, and consists in measuring the frequency dispersion of the rotation angle of the polarization ellipse and degree of ellipticity of polarization ellipse of the probe beam induced by the stronger optical field of the elliptically polarized pump beam by tuning its frequency ω [51]. Very cumbersome invariant expression for the nonlinear polarization $\mathbf{P}^{(3)}(\omega - \omega + \omega_1)$ can be obtained by direct summation in expression (9.19), taking account of (9.22) and (9.25). It is possible also investigate the dispersion of the rotation angle of the polarization ellipse and degree of ellipticity of polarization ellipse of the signal beam at frequency $2\omega - \omega_1$ (as in ordinary CARS on the cubic local nonlinearity $\chi_{ijkm}^{(3)}(2\omega - \omega_1; \omega, \omega, -\omega_1)$) induced due to nonlocal susceptibility $\gamma_{ijkml}^{(3)}(2\omega - \omega_1; \omega, \omega, -\omega_1)$ [52].

The effect of nonlinear optical activity, being understood as the intensity-dependent polarization plane rotation for the linearly polarized light in a medium with cubic nonlinearity, was predicted many years ago [8]. Initially it was explained solely by the spatial dispersion of the nonlinear optical response of the medium, described by nonlocal susceptibility $\gamma_{ijkml}^{(3)}(\omega + \omega - \omega; \omega, \omega, -\omega)$. At the earlier stage nonlinear optical activity was unjustifiably opposed to the self-rotation of the polarization ellipse [7]. The intensity-dependent rotation angle of the principal axis of the polarization ellipse increased with rising the ellipticity degree of the polarization ellipse of the incident plane wave and completely disappears for the linearly polarized light (zero degree of ellipticity). Both these effects, responsible for the polarization ellipse rotation and deformation take place in the medium with space group of symmetry $\infty\infty$ in a plane wave approximation, because the expression for $\mathbf{P}^{(3)}(\omega)$ can be written in this case in the following invariant form:

$$\begin{aligned} \mathbf{P}^{(3)}(\omega) = & 2\{(\chi_1^{(3)} + \chi_2^{(3)})\mathbf{E}(\omega)|\mathbf{E}(\omega)|^2 + \chi_2^{(3)}\mathbf{E}^*(\omega)(\mathbf{E}(\omega) \cdot \mathbf{E}(\omega)) + i\tilde{\gamma}_1^{(3)}\{|\mathbf{E}(\omega)|^2[\mathbf{k} \times \mathbf{E}(\omega)] \\ & + \mathbf{E}(\omega)(\mathbf{E}^*(\omega) \cdot [\mathbf{k} \times \mathbf{E}(\omega)])\} + i\tilde{\gamma}_2^{(3)}[\mathbf{E}^*(\omega) \times \mathbf{k}](\mathbf{E}(\omega) \cdot \mathbf{E}(\omega)). \end{aligned} \quad (9.28)$$

Here $\tilde{\gamma}_{1,2}^{(3)}$ are expressed through $\gamma_{1,2,3,4,5,6}^{(3)}(\omega, \omega, -\omega)$ and $\gamma_{1,2,3,4,5,6}^{(3)}(-\omega, \omega, \omega)$. Under weak dependence of $\gamma_{1,2,3,4,5,6}^{(3)}$ on frequencies one can assume $\tilde{\gamma}_2^{(3)}$ is approximately equal to zero.

Due to symmetry reason tensors $\chi_{ijk}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$, $\gamma_{ijklm}^{(3)}(\omega_1 + \omega_2 + \omega_3; \omega_1, \omega_2, \omega_3)$ and $\chi_{ijklm}^{(4)}(\omega_1 + \omega_2 + \omega_3 + \omega_4; \omega_1, \omega_2, \omega_3, \omega_4)$ of medium with space group of symmetry $\infty\infty m$ identically equal to zero and many of the above mentioned nonlinear optical effects do not occur in ordinary isotropic medium.

9.4 Polarization Singularities in the Electric Field at a Sum-Frequency Generated by Two Collinear Elliptically Polarized Beams in the Bulk of a Nonlinear Isotropic Medium with Spatial Dispersion of Nonlinear Optical Response

9.4.1 Formation of C-Type Polarization Singularities in the Beam Cross-Section at the Sum Frequency Generated by Uniformly Elliptically Polarized Gaussian Beams

The present section focuses on the study of formation of *L*-type and *C*-type polarization singularities in the beam cross section at the sum frequency generated in the bulk of an isotropic gyrotropic medium by uniformly collinear elliptically polarized Gaussian fundamental beams with frequencies $\omega_{1,2}$ propagating coincide with the z axis. The electric field strength of each of the beams is given by the formula (9.1). Polarization of radiation in the cross section of the signal-beam at sum frequency in this case [53] is determined only by the polar angle and is the same along any arbitrarily chosen radial direction defined by the vector, which originates from the beam center, where the intensity is zero. The cylindrical symmetry of the problem prohibits the appearance of *C*-points in the plane of the beam cross-section at frequency $\omega_1 + \omega_2$, but allows for the appearance of *C*-lines.

Papers [9, 53] describe in detail the approximations made and the method for finding the solution of the parabolic equation

$$\left(\frac{\partial}{\partial z} - \frac{i}{2k_{SF}} \Delta_{\perp} \right) E_{\pm}^{\text{SF}} = \frac{2\pi i k_{SF}}{\epsilon} P_{\pm}^{\text{SF}} \exp[i\Delta k(z - l_0)] \quad (9.29)$$

for the slowly varying envelopes E_{\pm}^{SF} of right- and left-hand circularly polarized vortex component of the signal wave at the sum frequency in an isotropic gyrotropic medium. In (9.29) k_{SF} is the wavenumber of the wave at frequency $\omega_1 + \omega_2$, $\Delta_{\perp} = \partial^2 / \partial x^2 + \partial^2 / \partial y^2$. If the medium has normal frequency dispersion, then the linear dielectric permittivity of the medium at the sum-frequency $\epsilon(\omega_1 + \omega_2) > \epsilon(\omega_{1,2})$, and $\Delta k = k_1 + k_2 - k_{SF} < 0$, where $k_{1,2}$ are the wavenumbers of the fundamental waves. In case of anomalous dispersion, the reversed

inequalities are valid. In (9.29) one neglect the linear gyration effects, as it was done in [53, 54], and assume wave-vector mismatch to be the same for both components E_{\pm}^{SF} . The right-hand side of this equation contains circularly polarized component P_{\pm}^{SF} of the vortex part $\mathbf{P}^{\text{SF}} = \mathbf{P}^{(2)} + (k_1 + k_2)^{-2} \vec{\nabla}_{\perp} (\vec{\nabla}_{\perp} \cdot \mathbf{P}^{(2)})$ of matter polarization $\mathbf{P}^{(2)}(\omega_1 + \omega_2)$, which appears as a result of the collinear propagation in quadratic medium of two uniformly elliptically polarized incident Gaussian fundamental beams (see (9.26)). The expressions for the electric field of these beams look like the (9.1). The symmetry axes of fundamental beams coincide with the z axis and the waist plane $z = l_0$ is common for them and located at a distance l_0 from the medium plane boundary. The initial conditions for $E_{\pm}^{\text{SF}}(x, y, z)$ are trivial and in approximation of small pump depletion (9.29) become linear heterogeneous ones. This allows us to write their solutions using the Green's function.

The circularly polarized electric field components at the sum-frequency $E_{\pm}^{\text{SF}}(r, \varphi, z, t)$ are expressed by the formulas [53]:

$$E_{\pm}^{\text{SF}}(r, \varphi, z, t) = F_0(r, z) E_{\pm}(\varphi) \exp[-i(\omega_1 + \omega_2)t + ik_{\text{SF}}(z - l_0)], \quad (9.30)$$

where $r\varphi z$ is the cylindrical coordinates. The function $F_0(r, z)$, whose explicit form is given in [9, 53], does not affect the degree of ellipticity $M(\varphi) = (|E_+^{\text{SF}}|^2 - |E_-^{\text{SF}}|^2) / (|E_+^{\text{SF}}|^2 + |E_-^{\text{SF}}|^2)$. In (9.30)

$$\begin{aligned} E_+(\varphi) = & 0.5 \left\{ \left[(1 + M_{01})^{1/2} (1 - M_{02})^{1/2} + (1 - k)(1 + M_{02})^{1/2} (1 - M_{01})^{1/2} \exp(2i\Psi) \right] \exp(i\varphi) \right. \\ & \left. + (1 + M_{01})^{1/2} (1 + M_{02})^{1/2} \exp(2i\Psi - i\varphi) \right\} \\ \equiv & A_+ \exp[i(\alpha_+ + \varphi)] + B_+ \exp[i(\beta_+ - \varphi)], \end{aligned} \quad (9.31)$$

$$\begin{aligned} E_-(\varphi) = & -0.5 \left\{ (1 - M_{01})^{1/2} (1 - M_{02})^{1/2} \exp(i\varphi) + [(1 - k)(1 + M_{01})^{1/2} (1 - M_{02})^{1/2} \right. \\ & \left. + k(1 + M_{02})^{1/2} (1 - M_{01})^{1/2} \exp(2i\Psi)] \exp(-i\varphi) \right\} \\ \equiv & A_- \exp[i(\alpha_- + \varphi)] + B_- \exp[i(\beta_- - \varphi)]. \end{aligned} \quad (9.32)$$

Here Ψ is the angle between the principal axes of the polarization ellipses of the fundamental waves at the medium input. The principal axis of the polarization ellipse of the wave with the frequency ω_1 is oriented along the straight line $\varphi = 0$. In (9.31) and (9.32) M_{01} and M_{02} are the degrees of ellipticity of these ellipses; $k = 1/(1 + k_2/k_1)$; A_{\pm} , B_{\pm} , α_{\pm} , and β_{\pm} are the real functions of k , $M_{01,02}$ and Ψ . One can see that the ellipticity degree of the wave at the frequency $\omega_1 + \omega_2$ is determined only by the polar angle φ and also has a period π . Therefore, the roots of the equation $M(\varphi) = 0$ set the angles, which define, in the plane of the signal beam cross section the directions of L -type lines. When the equation $M(\varphi) = 0$ has two

roots [9], in the cross of the beam at frequency $\omega_1 + \omega_2$ there are two L -lines intersecting at its center ($r=0$). These lines are the boundaries of the four sectors. The directions of rotation of the electric field vectors in two adjacent sectors are opposite. Singularity of the C -type polarization appears at those points of the beam cross section at the sum-frequency in which one of its circularly polarized components vanishes. One can see from (9.31), (9.32) that $E_+^{\text{SF}}=0$ at $A_+=B_+$ and $\varphi=(\beta_+-\alpha_+)/2+\pi/2+n\pi$, and $E_-^{\text{SF}}=0$ at $A_-=B_-$ and $\varphi=(\beta_--\alpha_-)/2+\pi/2+n\pi$, where $n=0, 1$. In other words, the C -line

$$\varphi=\varphi_+=\Psi+\frac{\pi}{2}-0.5 \arg[k(1+M_{01})^{1/2}(1-M_{02})^{1/2}+(1-k)(1+M_{02})^{1/2}(1-M_{01})^{1/2} \exp(2i\Psi)], \quad (9.33)$$

at the points of which the electric field vector rotates clockwise, appears when the condition

$$k^2(1+M_{01})(1-M_{02})+(1-k)^2(1-M_{01})(1+M_{02})+2k(1-k)(1-M_{01}^2)^{1/2}(1-M_{02}^2)^{1/2} \cos 2\Psi-(1+M_{01})(1+M_{02})=0 \quad (9.34)$$

is met, and the C -line

$$\varphi=\varphi_-=\frac{\pi}{2}+0.5 \arg\left[k \exp(2i\Psi)(1-M_{01})^{1/2}(1+M_{02})^{1/2}+(1-k)(1+M_{01})^{1/2}(1-M_{02})^{1/2}\right] \quad (9.35)$$

at the points of which the electric field vector rotates counter-clockwise, appears if

$$k^2(1-M_{01})(1+M_{02})+(1-k)^2(1+M_{01})(1-M_{02})+2k(1-k)(1-M_{01}^2)^{1/2} \times(1-M_{02}^2)^{1/2} \cos 2\Psi-(1-M_{01})(1-M_{02})=0. \quad (9.36)$$

Figure 9.4 shows in grey the domains of the ellipticity degree of the fundamental waves, for which L -type polarization singularities do not appear in the beam cross section at the sum frequency. For other values of M_{01} and M_{02} , there are two L -lines intersecting on the symmetry axis of the signal beam. At the points of solid curves bounding the gray areas (their equations are given in [9]), two L -lines merge into one. The dashed curves in Fig. 9.4 relate the values of M_{01} and M_{02} for which C -type lines appear in the beam cross section at the sum-frequency.

Figure 9.5 shows the typical distribution patterns of polarization ellipses in the plane of the beam cross section at the sum-frequency. The coordinates x and y are normalized by $w_3=w_1w_2(w_1^2+w_2^2)^{-1/2}$, where $w_{1,2}$ are fundamental radiation beams half-widths. The sum of the squares of the lengths of the semiaxes of each ellipse in this figure is proportional to the intensity at the beam point, which

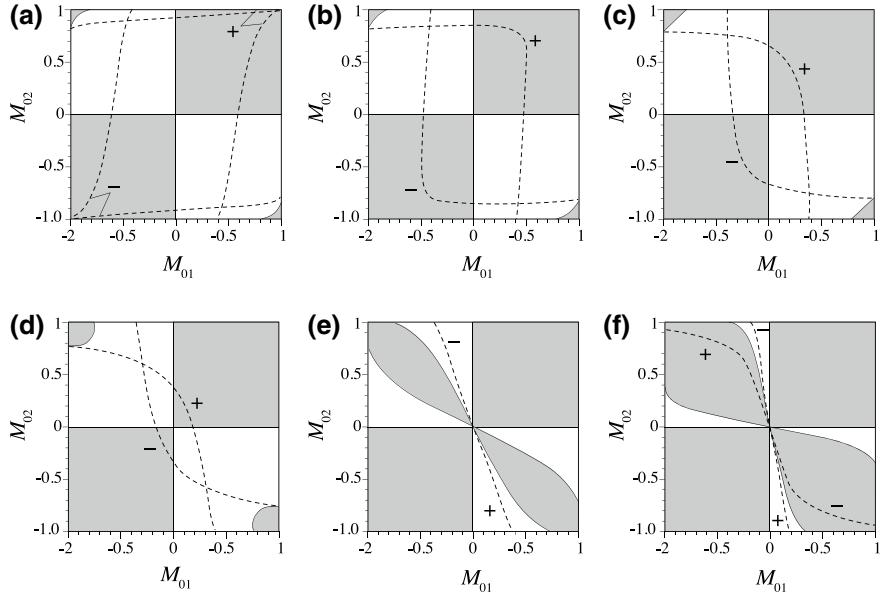


Fig. 9.4 (Adopted from [53]) Domains of the ellipticity degree of the fundamental waves corresponding to the generation of polarization singularities of C - and L -types in the beam cross section at the sum-frequency at $k_2 = 2k_1$ and $\Psi = \pi/2$ (a), $\pi/3$ (b), $\pi/4$ (c), and $\pi/6$ (d), as well as at $\Psi = 0$ and $k_2/k_1 = 2$ (e) and 5 (f)

coincides with its center; the ratio of the lengths of the semiaxes is uniquely expressed in terms of the degree of ellipticity of the wave at the same point in space; and the angle of inclination of its principal axis coincides with $\Phi(\varphi) = 0.5 \arg(E_+^{SF} E_-^{SF*})$ in the center of the ellipse. If the values of M_{01} and M_{02} correspond to the points of the dashed curve, which are within the grey area, then in the beam cross section at the sum-frequency there appears only one C -line (Fig. 9.5a), if to the points inside the white area there appear two L -lines and one C -line (Fig. 9.5b).

It is easy to see that $M(\varphi_{\pm}) = \mp 1$ and $|\Phi(\varphi_{\pm} + 0) - \Phi(\varphi_{\pm} - 0)| = \pi/2$, i.e., the rotation angle of the principal axis of the polarization ellipse at the points φ_{\pm} experiences a jump by $\pi/2$. This is illustrated in Fig. 9.6a, which shows the degree of ellipticity and rotation angle of the principal axis of the polarization ellipse versus the polar angle φ , where $0 \leq \varphi \leq \pi$. If $\Psi \neq 0$, then for the values of $M_{01,02}$, corresponding to the intersection point of the dashed curves (Fig. 9.4) constructed from equalities (9.33) and (9.35), respectively, there can appear two L -lines and two C -lines in beam cross section at the sum-frequency. The signal beam cross section is divided in this case by two C -lines into the four sectors, within each of which the degree of ellipticity changes from -1 to 1 and the rotation angle of the principal axis of the polarization ellipse remains unchanged.

Figure 9.6b shows one of these sectors and parts of two adjacent sectors to the right and left. As the Ψ decreases, the points of intersection of the dashed curves,

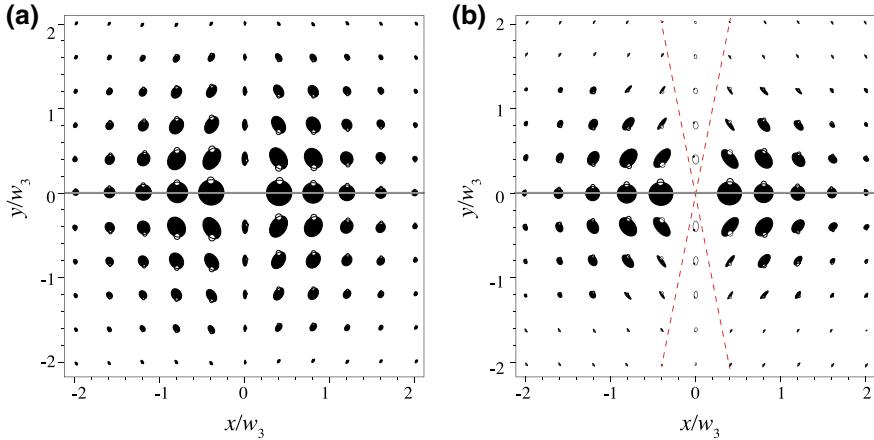


Fig. 9.5 (Adopted from [53]) Transverse spatial distribution of polarization in the beam at the sum-frequency at $M_{01} = -0.25$, $k_2 = 2k_1$ and $\Psi = \pi/2$ and $M_{02} = -0.938$ (a) and 0.903 (b). Solid lines show the C -lines, and dashed lines— L -lines

denoted by signs “plus” and “minus”, move to the point $M_{01,02} = 0$ in Fig. 9.4. In this case, the sizes of two of the four aforementioned sectors are reduced and the radiation intensity in smaller sectors drops substantially. In the limiting case $M_{01,02} = 0$, “small” sectors collapse into a line on which the radiation intensity is zero. The appearance of two C -lines is impossible, if $\Psi = 0$.

9.4.2 *Polarization Singularities in a Sum-Frequency Light Beam Generated by a Bichromatic Singular Beam*

In the present section we analyze the creation, evolution, and interaction of polarization singularities in a sum-frequency beam generated in an isotropic chiral medium by bichromatic singular beam [55]. We call so the superposition of regular elliptically polarized Gaussian beam \mathbf{E}_1 at frequency ω_1 , which is determined by (9.1), and circularly polarized Laguerre-Gaussian beam

$$\mathbf{E}_2(x, y, z, t) = E_{20} \left[\mathbf{e}_+ + \frac{i\mathbf{e}_z(\mathbf{e} \cdot \nabla_\perp)}{k_2} \right] \frac{(x + imy)}{w_2 \beta_2^2(z)} \exp \left(-\frac{x^2 + y^2}{w_2^2 \beta_2^2(z)} - i\omega_2 t + ik_2(z - l_0) \right). \quad (9.37)$$

Here E_{20} is the amplitude, w_2 is the right-hand circularly polarized beam half width, $\beta_2(z) = 1 + 2i(z - l_0)/(k_2 w_2^2)$, $m = \pm 1$ is the charge of the Laguerre-Gaussian mode (the case of left-sided circular polarization can be considered in a similar way).

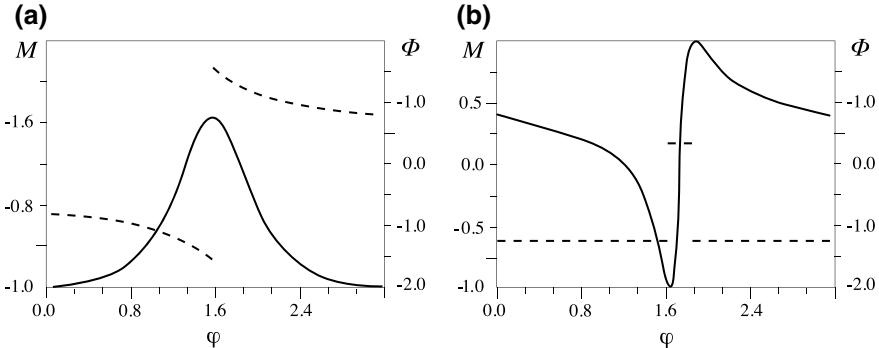


Fig. 9.6 (Adopted from [53]) Dependences of the degree of ellipticity (solid curves) and the rotation angle of the principal axis of the polarization ellipse (dashed curves) on the polar angle in the beam cross section at the sum-frequency at $k_2=2k_1$ and $M_{01}=-0.25$, $M_{02}=-0.938$, $\Psi=\pi/2$ (a), $M_{01}=-0.3$, $M_{02}=0.6$, $\Psi=\pi/6$ (b)

Firstly one finds the nonlinear polarization vector $\mathbf{P}^{(2)}(\omega_1 + \omega_2)$ substituting (9.1) and (9.37) in (9.26). After that it is necessary to separate its solenoidal part \mathbf{P}^{SF} (see Sect. 9.4.1) and substitute it in (9.29). As in Sect. 9.4 we neglect the linear gyration and assume wave vector mismatch to be the same for both components E_{\pm}^{SF} . The initial conditions are trivial, and in approximation of small pump depletion (9.29) one can write their solutions using the Green's function in quadratures. Each components of E_{\pm}^{SF} consists of a “central core” with a Gaussian-like intensity profile and a “frame,” similar to Laguerre-Gaussian modes of second order. If the diffraction lengths of two fundamental beams are equal, then the signal beam does not contain any polarization singularities.

The search for C-points in the cross section of signal beam and the reconstruction of C-lines were done by the following numerical algorithm [55]. At first, the grid values of integrals were found numerically for the consequence of the propagation coordinate values. After that the singularity patterns were reconstructed using piecewise interpolation methods. Polarization singularities emerge in the signal beam if the vector field \mathbf{P}^{SF} itself has polarization singularities at the medium's border. From now on we will call the latter singularities *the generators* or *G-points* (left-hand *G-points* if $P_+^{\text{SF}}=0$ and right-hand generators if $P_-^{\text{SF}}=0$). The amount and characteristics of generators and their positions depend on parameters of the beams (9.1) and (9.37). Their trajectories form *G-lines*. Polarization singularities in the signal beam have the same topological charges as the corresponding generators and are located close to them in the cross sections which are close enough to the medium's border. It is easy to prove that the distribution of *G-points* in the transversal plane $z=\text{const}$ has the center of symmetry, and their amount cannot exceed four for each handedness of rotation of the \mathbf{P}^{SF} vector. Symmetrically positioned *G-points* have identical topological charges.

Varying the waists of the beams $w_{1,2}$ and ellipticity degree M_0 (assuming the wave-vector ratio k_2/k_1 to be constant), one can create one of the three possible

types of configurations of the medium's nonlinear polarization field. The first type is realized with $m\{M_0 - M^*\} < 0$, where $M^* = [k_2^2 - (k_1 + k_2)^2]/[k_2^2 + (k_1 + k_2)^2]$. The beam of nonlinear polarization contains one pair of left-hand G -points in the $z=l_0$ plane with total topological charge equal to m . These generators also exist at any other z (Fig. 9.7a, b). This configuration of vector field \mathbf{P}^{SF} generates an analogous polarization structure of the electric field at the sum frequency. However, unlike G -lines each C -line has a helical structure if the dimensionless wave-vector mismatch $\Delta k \neq 0$ and the mean "step" of the spiral decreases as the absolute value $|\Delta k|$ grows. There also can be a more complicated transformation resulting in such a curving of C -lines that they start to intersect some of the transversal planes more than once. This transformation leads to the appearance of additional C -points' birth and annihilation events. Emerging two C -points have opposite topological charges and the helicity of each spiral is also opposite.

Figure 9.8 illustrates C -lines (curved lines) in the sum-frequency beam and G -lines (straight lines) of vector field \mathbf{P}^{SF} for different values of the wave-vector mismatch. White and black stars denote the points of space where pairwise creation (white stars) or annihilation (black stars) of C -points takes place. Red (solid) lines in (a, c) and markers in (b, d) denote the singularities with positive (1/2) topological charge and blue (dashed) lines and white markers denote ones with negative (-1/2) charge. We notice that each star connects two cuts of C -lines with opposite topological charges.

The second type of the nonlinear medium's polarization field configuration is realized when both $m\{M_0 - M^*\} > 0$, and $k_1 w_1^2 > k_2 w_2^2$ are valid. In this case there are two pairs of G -points in the $z=l_0$ plane with zero total topological charge. The generators exist only in the planes $z=z_1$, if z_1 is in the interval $[-z_2, z_2]$. In the plane $z=l_0 - z_2$ the birth of two pairs of G -points takes place, and in $z=l_0 + z_2$ their pairwise annihilation occurs (Fig. 9.7c, d). The value $z_2 = 0.5(k_1 k_2)^{1/2} w_1 w_2 [1 + \lambda_1 - (\lambda_1^2 + 2\lambda_1)^{1/2}]^{1/2}$, where $\lambda_1 = (M_0 - M^*) \xi (1 + M^*)^{-1} (1 - M_0)^{-1}$, if $m=1$, and $\lambda_1 = (M^* - M_0) \xi (1 - M^*)^{-1} (1 + M_0)^{-1}$ if $m=-1$, where $\xi = (k_1 w_1^2 + k_2 w_2^2)^2 / k_1 k_2 w_1^2 w_2^2$.

Four corresponding cuts of G -lines start and end in places of birth and annihilation of G -points ($z=l_0 \mp z_2$), forming a closed loop. If $l_0 > z_2$, then the fundamental beams are focused in such a way that the whole loop is located inside the medium, where $z > 0$. As in the configuration of the first type, when $\Delta k = 0$, C -lines in the sum-frequency beam have almost the same structure as the G -lines do. The presence of wave-vector mismatch in this case leads to the deformation of the C -lines loop, and the growth of $\Delta k l_3$ causes the formation of additional pairs of C -points and appearance of new loops of C -lines which are not connected with the primal loop. Here $l_3 = 0.5(k_1 + k_2)w_3^2$, $w_3^2 = (w_1 w_2)^2 / (w_1^2 + w_2^2)$. Further increasing of the absolute value of $\Delta k l_3$ leads to the splitting of the primal loop into two new loops, and the cuts of C -lines become helical. The loops of left-hand C -lines for different values of $\Delta k l_3$ and fixed parameters of the fundamental beams are shown in Fig. 9.9.

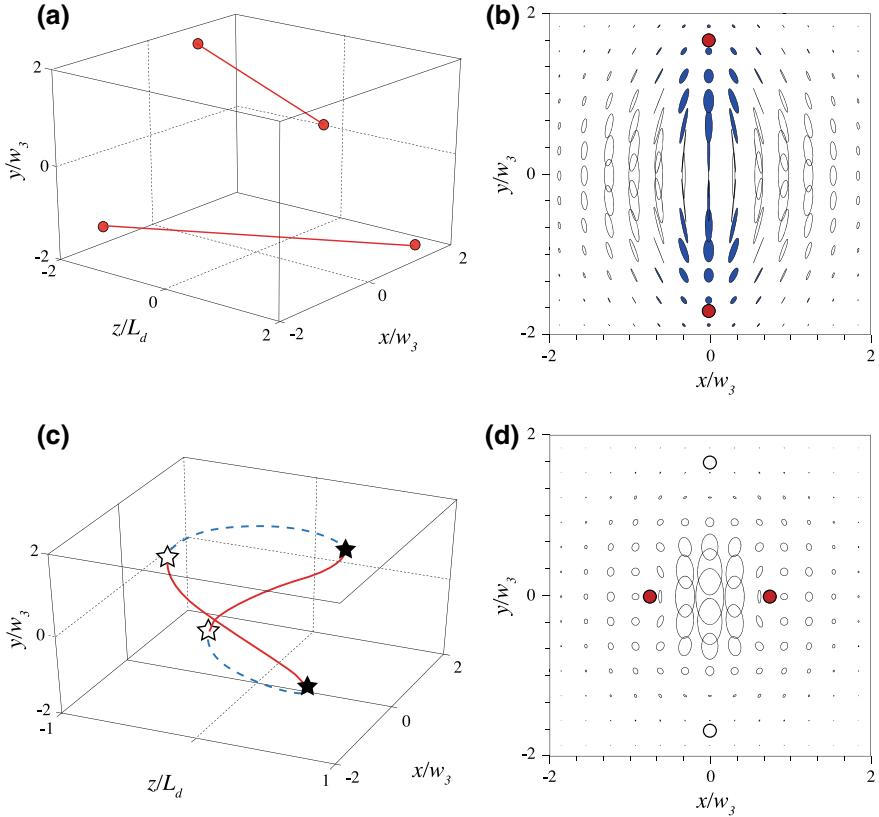


Fig. 9.7 (Adopted from [55]) Left-hand G -lines (**a**, **c**) and transversal distributions of the nonlinear medium polarization field at the position of the waist of the fundamental beams (**b**, **d**) for the first (**a**, **b**) and the second (**c**, **d**) types of vector field \mathbf{P}^{SF} configurations. The charge $m=1$, and beams' parameters are $w_2/w_1=2$, $k_2/k_1=1.4$, $M_0=-0.7$ (**a**, **b**) and $w_2/w_1=0.3$, $k_2/k_1=2$, $M_0=0$ (**c**, **d**). The plane $z=0$ in the figure corresponds to the position of the fundamental beams' waists. Red (full) and blue (dashed) lines correspond to C -lines with positive and negative topological charges, respectively. C -points with positive and negative charges are marked by red (filled) and white circles. Blue (filled) ellipses are left-hand polarized and empty ellipses are right-hand polarized

Finally, the third type of configuration of nonlinear polarization field \mathbf{P}^{SF} is realized when both $m\{M_0 - M^*\} > 0$ and $k_1 w_1^2 < k_2 w_2^2$ are valid. The beam of this type does not contain singularities in the waist plane $z=l_0$, although two pairs of G -points exist in the planes $z=z_1$, where z_1 satisfies the following relations: $z_1 < l_0 - z_3$ or $z_1 > l_0 + z_3$. In this case total topological charge of G -points is zero as in the second type of \mathbf{P}^{SF} field configuration. In the $z=l_0 + z_3$ plane birth of G -points takes place, and in $z=l_0 - z_3$ their pairwise annihilation does. The characteristic coordinate $z_3 = (l_1 l_2)^{1/2} [1 + \lambda_1 + (\lambda_1^2 + 2\lambda_1)^{1/2}]^{1/2}$, is found in a similar

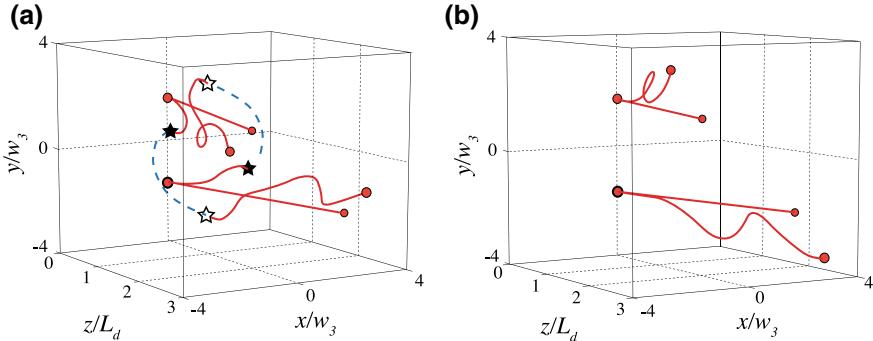


Fig. 9.8 (Adopted from [55]) Left-hand C -lines (curved lines) and G -line (straight lines) for $\Delta k l_3 = -5$ (a), $\Delta k l_3 = 5$ (b), and $w_2/w_1 = 2$, $k_2/k_1 = 1.4$, $m = 1$, $M_0 = -0.7$, $l_0 = 0$. In figure (a) each C -line forms a left-hand screw, and in figure (b) it forms a right-hand one. Red (full) and blue (dashed) lines correspond to C -lines with positive and negative topological charges, respectively. C -points with positive and negative charges are marked by red (filled) and white circles

way as z_2 . The value of z_3 cannot be less than the minimum of two diffraction lengths $l_{1,2} = k_{1,2} w_{1,2}^2 / 2$ of the fundamental beams. Despite the fact that C -points' formation induced by G -points is still possible in this type of configuration, it is not of practical interest, because \mathbf{P}^{SF} tends to zero as $|z - l_0|$ increases and the efficiency of the sum-frequency generation is relatively small.

In the present section only the circular polarization singularities are considered in sum frequency generation process. Following the introduced “generator singularity” concept, one can show that L -surfaces appear in the signal beam as well. However, detailed analysis of these singularities is more complicated compared to C -lines, because L -surfaces are objects of higher dimension and they are harder to be captured by numerical methods.

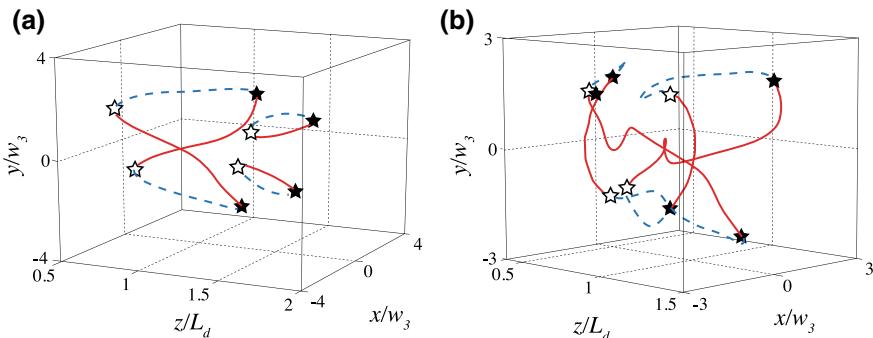


Fig. 9.9 (Adopted from [55]) Left-hand C lines in the sum-frequency beam for $\Delta k l_3 = -4$ (a), $\Delta k l_3 = -8$ (b), and $w_2/w_1 = 0.3$, $k_2/k_1 = 2$, $m = 1$, $M_0 = 0$, $l_0 = l_3$. C -lines form three separate loops in figure (a) and two loops in figure (b). Red (full) and blue (dashed) lines correspond to C -lines with positive and negative topological charges, respectively

9.5 Singularities of the Sum-Frequency Light Field Polarization Arising upon Reflection on Incident Elliptically Polarized Gaussian Beams from the Surface of a Isotropic Medium with Spatial Dispersion of Nonlinear Optical Response

9.5.1 Nonuniform Transverse Distribution of the Light Intensity and Polarization upon Sum-Frequency Generation from the Surface

A reflected beam at the sum frequency $\omega_3 = \omega_1 + \omega_2$ emerges due to nonlinear optical responses of the surface of a medium and its volume caused by monochromatic fundamental beams at frequencies $\omega_{1,2}$. The first of them is related to the difference in the symmetry of the surface layer of the isotropic gyroscopic medium (symmetry group ∞) from the symmetry of its thickness (symmetry group $\infty\infty$). We consider it by using the boundary conditions (9.16) for the electromagnetic field. The reflected sum-frequency signal emerged by the thin surface layer is related to the second term in the expression (9.17) for surface current density of coupled charges, which can be written in the form:

$$i_{ikl}(\omega_3) = \kappa_{ikl}^{(2)}(\omega_3; \omega_1, \omega_2) E_k^{(v)}(\omega_1) E_l^{(v)}(\omega_2). \quad (9.38)$$

To find the nonlinear optical response of the medium thickness caused by monochromatic waves at frequencies $\omega_{1,2}$ we use (9.18). Consider weakly diverging monochromatic fundamental beams (9.1), which fall collinearly on the surface of the isotropic gyroscopic medium. Its common symmetry axis coincides with the z axes. By assuming that the characteristic changes in the electric field strength in the cross-section planes of incident, reflected and refracted beams occur at distances much larger than wavelengths, one can represent $E^{(v)}(\omega_{1,2,3})$ and $E^{(t)}(\omega_{1,2,3})$ at medium interface $z=0$ by using the spatial Fourier integral, in the form of superposition of monochromatic plane waves with different components of the wave vector lying in plane $z=0$. To simplify the problem we will assume below that absorption is absent in the medium and all the spatial Fourier components of electric field in the medium have real wave vectors. Taking into account the smallness of the angular divergence of the fundamental beams (9.1), in determining the field we will allow only for linear terms in the components of the wave vectors of the fundamental beams lying in the plane perpendicular to the z axes, and in exponents of Fourier integrals (where they are absent) – quadratic terms. Spatial Fourier components of electric fields $E^{(v)}$ and $E^{(t)}$ at medium surface can be readily expressed via spatial Fourier transforms of the electric fields of incident waves at frequencies $\omega_{1,2}$ (it is easy to obtain them from (9.1)) using Fresnel formulas. As a result, we obtain the relations for spatial Fourier components of electric fields

$\mathbf{E}^{(v)}(\omega_{1,2})$ on the medium surface. Knowing the latter it is possible to find the expressions for the spatial Fourier components of medium polarization and surface current density of coupled charges at frequency $\omega_1 + \omega_2$. To calculate the spatial Fourier components of the electric field strength of the reflected wave at the sum frequency, we substitute into boundary conditions spatial Fourier components of the surface current density of coupled charges and the electric field induction simply connected with the medium polarization and then to solve the derived system of algebraic equations. As a result, one obtain expressions for spatial Fourier components constituting the electric field strengths of the reflected sum-frequency wave near the medium surface. Finally, by performing integration, it is possible to find the intensity and polarization distributions in the plane of the reflected beam cross-section at the sum frequency near the medium surface. The expressions for right- and left-hand circularly polarized component E_{\pm}^{SF} of the signal wave at the frequency $\omega_1 + \omega_2$ take the following form [56]:

$$\begin{aligned} E_{\pm}^{\text{SF}}(r, \varphi, z) = & -D(r, z) \left\{ \exp(\pm i\varphi) [(C_{0\pm} - C_{2\pm})(1 \pm M_{01})^{1/2} (1 \mp M_{02})^{1/2} \right. \\ & \times \exp[\pm i(\Psi_1 - \Psi_2)] + (C_{0\mp} - C_{1\pm})(1 \mp M_{01})^{1/2} (1 \pm M_{02})^{1/2} \exp[\mp i(\Psi_1 - \Psi_2)] \\ & \left. - \exp(\mp i\varphi)(C_{1\pm} + C_{2\pm})(1 \pm M_{01})^{1/2} (1 \pm M_{02})^{1/2} \exp[\pm i(\Psi_1 + \Psi_2)] \right\}, \end{aligned} \quad (9.39)$$

in the cylindrical coordinates $r\varphi z$. Here $E_{01,02}$ are the amplitudes of fundamental homogeneously polarized beams, $M_{01,02}$ are the degrees of ellipticity of polarization ellipses, $\Psi_{01,02}$ are the angles determined their orientations,

$$D(r, z) = \frac{8\sqrt{2}\pi i E_{01} E_{02}}{(1+n_1)(1+n_2)(1+n_3)\omega_3 w_3^2 \beta_3^2(z)} \frac{r}{\beta_3(z)} \exp \left[-\frac{r^2}{w_3^2 \beta_3(z)} - i(k_3 z + \omega_3 t) \right], \quad (9.40)$$

where $\beta_3(z) = 1 - iz/l_3$, $n_{1,2,3} = \sqrt{e(\omega_{1,2,3})}$. The coefficients $C_{0\pm}$, $C_{1\pm}$ and $C_{2\pm}$ in expression (9.39) depend on the medium parameters, $\omega_{1,2,3}$ and have the form

$$\begin{aligned} C_{0\pm} = & n_3 b_1 \pm i(n_3 b_7 + \xi_{\omega} c \chi^{(2)}(\omega_1, \omega_2)/n_3) \\ & + i\xi_{\omega} [(w_2^2 \gamma_1 + w_1^2 \gamma_2)/(w_2^2 + w_1^2) + (n_1 \omega_1 \gamma_1 + n_2 \omega_2 \gamma_2)/n_3 \omega_3], \end{aligned} \quad (9.41)$$

$$C_{1\pm} = (l_3/l_1) \left\{ n_1 b_4 \pm i(n_1 b_6 + \xi_{\omega} c \chi^{(2)}(\omega_1, \omega_2)/n_1) + i\xi_{\omega} \gamma_4 n_2 \omega_2 (n_2 l_2 - n_1 l_1)/n_1 \omega_3 n_2 l_2 \right\}, \quad (9.42)$$

$$C_{2\pm} = (l_3/l_2) \left\{ n_2 b_3 \mp i[n_2 b_5 + \xi_{\omega} c \chi^{(2)}(\omega_1, \omega_2)/n_2] - i\xi_{\omega} \gamma_3 n_1 \omega_1 (n_2 l_2 - n_1 l_1)/n_2 \omega_3 n_1 l_1 \right\}, \quad (9.43)$$

where $\xi_\omega = \omega_3 / (n_1\omega_1 + n_2\omega_2 + n_3\omega_3)$. Expressions (9.41)–(9.43) contain six out of seven independent tensor components of the quadratic response of the medium surface $\kappa_{ikl}^{(2)}(\omega_3; \omega_1, \omega_2)$: $b_1 = \kappa_{zxx}^{(2)} = \kappa_{zyy}^{(2)}$, $b_2 = \kappa_{zzz}^{(2)}$, $b_3 = \kappa_{yyz}^{(2)} = \kappa_{xxz}^{(2)}$, $b_4 = \kappa_{xzx}^{(2)} = \kappa_{zyy}^{(2)}$, $b_5 = \kappa_{xyz}^{(2)} = -\kappa_{yxz}^{(2)}$, $b_6 = \kappa_{yzx}^{(2)} = -\kappa_{xzy}^{(2)}$, $b_7 = \kappa_{zxy}^{(2)} = -\kappa_{zyx}^{(2)}$. These components are specified in the crystal-physical coordinate system coinciding with xyz ($x = r\cos\varphi, y = r\sin\varphi$), because the medium itself and its surface layer are isotropic, an arbitrary choice of directions of x and y axes is possible. The same expressions include components of tensors $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2) \equiv \gamma_{ijkl}^{(2)1}$ and $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_2, \omega_1) \equiv \gamma_{ijkl}^{(2)2}$, which characterize the spatial dispersion of the quadratic optical response of the medium volume: $\gamma_{1,2} = \omega_3 \gamma_{xxy}^{(2)1,2}$, $\gamma_3 = \omega_3 \gamma_{xyx}^{(2)1}$ and $\gamma_4 = \omega_3 \gamma_{xyx}^{(2)2}$ (see (9.18) and (9.24)). Let us emphasize that expressions (9.39)–(9.43) also contain terms found in the first approximation with respect to the parameters $\lambda_{1,2,3}/w_{1,2,3}$.

Expression (9.39) allows one to calculate the intensity, the degree of ellipticity M_3 and the rotation angle Ψ_3 of the principal axes of the polarization ellipse of radiation at the sum frequency. If $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2) = \gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_2, \omega_1) = 0$ and polarizations of incident waves are linear, the sum frequency signal is also linearly polarized at all the points of the reflected beam cross section. In the case of equal degrees of ellipticity of incident beams and $n_1 l_1 = n_2 l_2$ the sum frequency signal is independent of the tensor $\chi_{ijk}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$ at equal orientations of principal axes of polarization ellipses of incident waves ($\Psi_{01} = \Psi_{02}$). If the principal axes of these ellipses are mutually perpendicular, the sum-frequency field is independent of the components of tensors $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_1, \omega_2)$ and $\gamma_{ijkl}^{(2)}(\omega_1 + \omega_2; \omega_2, \omega_1)$. In the case of the beams incident on the medium are circularly polarized in opposite directions ($M_{01} = -M_{02} = \pm 1$), then $M_3 = \tilde{M}_3 = \text{const}$ and Ψ_3 differs from φ by the constant: The electric field vector of radiation at the sum frequency rotates counterclockwise along the ellipses shown in Fig. 9.10a. The electromagnetic field distribution in this case can change rather strongly with varying ω_1/ω_2 and w_1/w_2 , which is shown in Fig. 9.11.

Figure 9.11 presents the dependences of \tilde{M}_3 on ω_1/ω_2 and w_1/w_2 . One can see that with increasing ω_1/ω_2 , \tilde{M}_3 can change from -1 to 1 . If $\hat{\kappa}^{(2)}(\omega_3) = 0$ and $n_1 l_1 = n_2 l_2$, $\tilde{M}_3 = 0$ and $\Psi_3 = \varphi$. In this case, the sum-frequency beam is radially polarized. Analysis of the dependence $M_3(\varphi)$ shows that the equation $M_3(\varphi) = 0$, defining the position of L -line in the plane of the reflected beam cross section, is quadratic with respect to $\tan \varphi$. Figure 9.10b shows the intensity and polarization distribution corresponding to this case. One can clearly see four sectors obtained due to intersection of straight lines at the center of the beam. Radiation within each of them is elliptically polarized and at the boundaries—linearly polarized. In this case, the rotation directions of the electric field vectors at the points lying in adjacent sectors are different (white and black ellipses).

9.5.2 *Polarization Singularities in Second Harmonic Beam Arising upon Reflection of Elliptically Polarized Gaussian Beam in the Case of Normal Incidence of Light*

One of the effective methods for studying nonlinear optically active media and thin films is based on the second harmonic generation from their surface. The conditions for second-harmonic generation and the technique for detecting the effects are well known and described in the literature [57–62]. A relatively long time ago some theoretical studies were carried out within the plane-wave approximation, where the influence of the spatial dispersion of the nonlinear optical response of a chiral medium [57–60] and the surface inhomogeneity of its optical properties [63] were taken into account in different ways. In some experiments there were attempts to separate the contributions from the surface and volume of the investigated material to the second-harmonic signal [64] and select the second-harmonic component emerging due to the gyrotropy of the medium [65, 66]. The authors of [61, 62] were the first to take into account the boundedness of the incident fundamental beam when calculating the intensity and polarization parameters of the second-harmonic signal from the surface of a medium with a spatial dispersion of quadratic nonlinearity.

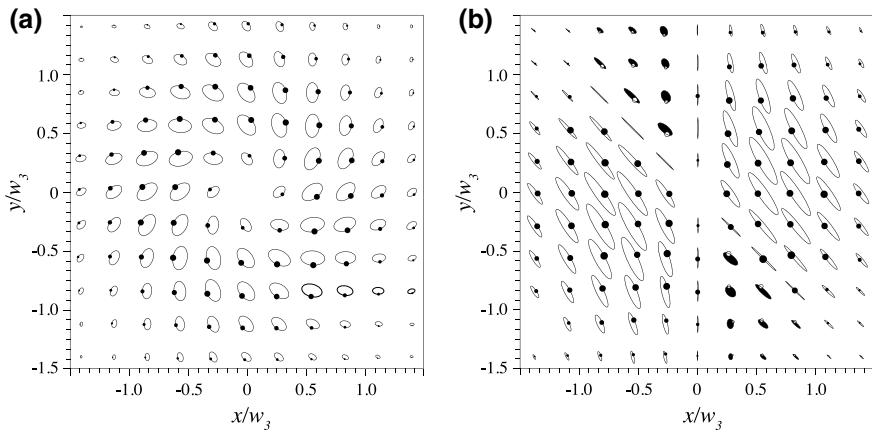


Fig. 9.10 (Adopted from [56]) Intensity and polarization distribution in a sum-frequency beam for the following parameters of incident radiation and nonlinear medium: $M_{01} = 1$, $M_{02} = -1$, $b_1 = b_3/3 = b_4 = b_5 = -b_6/2 = 2b_7$, $c\chi^{(2)}(\omega_1, \omega_2) = b_1$, $w_1/w_2 = 2$ (a) and $M_{01} = M_{02} = 0$, $\Psi_{01} = \Psi_{02} = 0$, $b_1 = b_3 = b_4/2 = b_5 = -b_6/2 = 2b_7$, $\chi^{(2)}(\omega_1, \omega_2) = 0$, $(w_1/w_1)^2 = n_2\omega_2/n_1\omega_1 = 1.548$ (b); $n_1 = 1.26$, $n_2 = 1.3$, $n_3 = 1.34$, $\omega_1 = 0.4\omega_3$, $\gamma_1 = 2\gamma_2/3 = 5\gamma_3/6 = 4\gamma_4/5 = b_1$

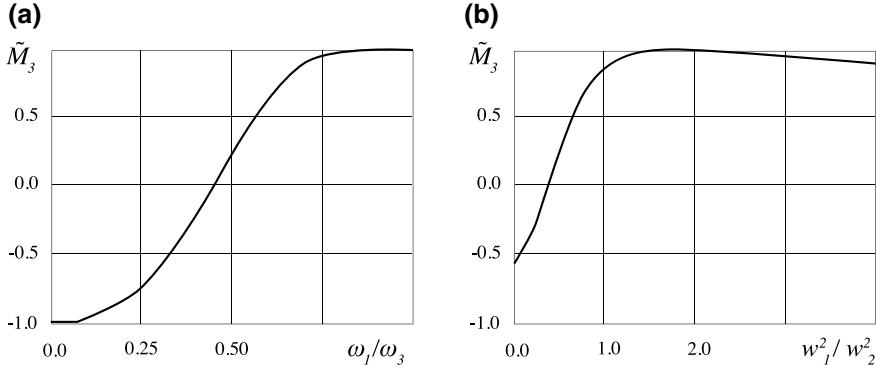


Fig. 9.11 (Adopted from [56]) Dependences of \tilde{M}_3 on ω_1/ω_3 at $w_1/w_2=0.5$ **(a)** and \tilde{M}_3 on w_1/w_2 at $\omega_1/\omega_3=0.4$ **(b)**; $M_{01}=1$, $M_{02}=-1$, $b_1=b_3/3=b_4=b_5=-b_6/2=2b_7$, $\gamma_1=2\gamma_2/3=5\gamma_3/6=4\gamma_4/5=b_1$, $c\chi^{(2)}(\omega_1, \omega_2)=b_1$, $n_1=1.26$, $n_2=1.3$, $n_3=1.34$

The expressions for right- and left-hand circularly polarized component E_{\pm}^{SH} of the signal wave at the double frequency one can obtain assumed that $\omega_1=\omega_2$ in formulae of previous section:

$$\begin{aligned} E_{\pm}^{\text{SH}}(r, \varphi, z) = & \frac{32\pi i E_0^2}{[1+n(\omega)]^2 [1+n(2\omega)] w^2 \omega [\beta(z)]^2} \left\{ -C_0 (1-M_0^2)^{1/2} \exp(\pm i\varphi) \right. \\ & + \left[(1+M_0)^{1/2} \times \exp(-i\varphi) + (1-M_0)^{1/2} \exp(i\varphi) \right] (1 \pm M_0)^{1/2} C_{\pm} \} \\ & \times \exp \left[-\frac{2r^2}{w^2 \beta(z)} - 2i\omega \left(t + \frac{z}{c} \right) \right]. \end{aligned} \quad (9.44)$$

Here E_0 is the amplitude of fundamental homogeneously polarized beam, M_0 is the degrees of ellipticity of polarization ellipse, $\beta(z)=1-2iz/(kw)^2$, $C_0=n(2\omega)b_1+i\omega\gamma_{xxxy}^{(2)}(2\omega; \omega, \omega)/n(2\omega)$, $C_{\pm}=n(\omega)(b_3 \pm ib_5)$, $b_1=\kappa_{zx}^{(2)}=\kappa_{zy}^{(2)}=\kappa_{yz}^{(2)}=\kappa_{xz}^{(2)}$, $b_5=\kappa_{xy}^{(2)}=-\kappa_{yx}^{(2)}=\kappa_{xy}^{(2)}=-\kappa_{yx}^{(2)}$. The condition for the occurrence of the L -line $\varphi=\varphi_L$, which makes the angle φ_L with the x axis, in the cross section of the second harmonic beam is $|E_+^{\text{SH}}(r, \varphi_L, z)|=|E_-^{\text{SH}}(r, \varphi_L, z)|$. It is shown [67] that, at certain values of the nonlinear-medium parameters and the incident light ellipticity, the last equation may have two roots,

$$\varphi_{L1, L2} = \arctan \left(-B/A \pm \sqrt{(B/A)^2 - C/A} \right), \quad (9.45)$$

if the values

$$A = \left[1 - (1 - M_0^2)^{1/2} \right] \left[M_0 n(\omega) (b_3^2 + b_5^2) + b_5 \omega \gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega) \right] - M_0 \left[M_0 b_5 \omega \gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega) + (1 - M_0^2)^{1/2} n(2\omega) b_1 b_3 \right], \quad (9.46)$$

$$B = (1 - M_0^2)^{1/2} \left[b_3 \omega \gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega) + n(2\omega) b_1 b_3 M_0 \right], \quad (9.47)$$

$$C = M_0 n(\omega) (b_3^2 + b_5^2) + (1 - M_0^2) b_5 \omega \gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega), \quad (9.48)$$

which enter into (9.45), satisfy the condition $B^2 > AC$. In this case, the cross section plane of the second-harmonic beam is divided by two L -lines ($\varphi = \varphi_{L1}$ and $\varphi = \varphi_{L2}$), which are intersected at the center of beam into four sectors, so that the polarization rotation direction changes every time when passing through the L -line. At $B^2 = AC$ we have $\varphi_{L1} = \varphi_{L2}$, and there is one L -line in the cross-section of the reflected beam and the polarization rotation is the same everywhere. If $B^2 < AC$, (9.45) has no solutions. Polarization singularities of the L -type do not arise in this case.

The conditions for the occurrence of C -line $\varphi = \varphi_{C+}$ ($\varphi = \varphi_{C-}$) with left-handed (right-handed) rotation of the electric field strength vector, can be written, respectively, as $E_+^{\text{SH}}(r, \varphi_{C+}, z) = 0$ ($E_-^{\text{SH}}(r, \varphi_{C-}, z) = 0$). Actually, each of these equations is a system of two equations, because the quantities E_{\pm}^{SH} are complex. Therefore, their solutions $\varphi = \varphi_{C+}$ and $\varphi = \varphi_{C-}$, exist only under some limitation on the parameters M_0 , b_1 , b_3 , b_5 and γ_0 . It is reasonable to introduce this limitation as a dependence of the polarization of incident radiation on the parameters of nonlinear chiral medium. At $M_0 = M_{0\pm} = \pm(G_{\pm} - G_1)/(G_{\pm} + G_1)$, where $G_1 = n^2(\omega)(b_3^2 + b_5^2)$ and $G_{\pm} = [\omega \gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega) \pm n(\omega)b_5]^2 + [n(\omega)b_3 - n(2\omega)b_1]^2$, the solutions to $E_+^{\text{SH}}(r, \varphi_{C+}, z) = 0$ and $E_-^{\text{SH}}(r, \varphi_{C-}, z) = 0$ take, respectively, the form

$$\varphi = \varphi_{C\pm} = \arctan \frac{n(\omega)b_3\sqrt{G_{\pm}} + [-n(2\omega)b_1 + n(\omega)b_3]\sqrt{G_1}}{n(\omega)b_5\sqrt{G_{\pm}} - [\pm\omega\gamma_{xxy}^{(2)}(2\omega; \omega, \omega) / n(2\omega) + n(\omega)b_5]\sqrt{G_1}}. \quad (9.49)$$

Having passed through the line $\varphi = \varphi_{C+}$ or $\varphi = \varphi_{C-}$ in the beam cross-section plane, the angle of rotation of the major axis of the polarization ellipse changes stepwise by $\pi/2$. Note that both values of the ellipticity $M_{0\pm}$ exist at any real values of b_1 , b_3 , b_5 and γ_0 . Only when all components of the tensors $\hat{\gamma}^{(2)}(2\omega; \omega, \omega)$ and $\hat{\kappa}^{(2)}(2\omega; \omega, \omega)$ are zero (the second harmonic signal is absent), $G_{\pm} + G_1 = 0$.

For example, Fig. 9.12 shows the polarization distributions in the cross section of the reflected signal at the double frequency. The ellipses plotted in the different areas of the figure, with centers at the coordinates (x_0, y_0) , are similar to the light polarization ellipses at the points with the same coordinates in the second harmonic beam cross section. The angle between the major axis of the ellipse centered at the point (x_0, y_0) and the x axis in Fig. 9.12 coincides with the inclination angle of the

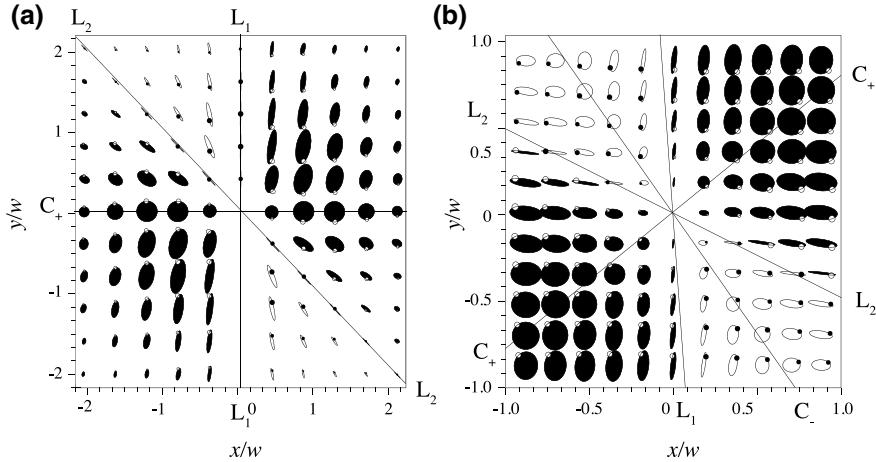


Fig. 9.12 (Adopted from [67]) Nonuniform distribution of light polarization in the second harmonic reflected beam at **a** $M_0=0$, $n(\omega)=1.33$, $n(\omega)=1.35$, $b_3/b_1=0.508$, $b_5=b_3$, $\omega\gamma_{xxy}^{(2)}(2\omega;\omega,\omega)/b_1=-1.823$ (one C -line and two L -lines) and **b** $M_0=-0.27$, $n(\omega)=1.32$, $n(2\omega)=1.34$, $b_3/b_1=1.1$, $b_5=0.322$, $\omega\gamma_{xxy}^{(2)}(2\omega;\omega,\omega)/b_1=-2.1$ (two C -lines and two L -lines). The ends of the C -and L -lines are marked by symbols L_1 , L_2 , C_+ and C_-

major axis of the polarization ellipse of second harmonic signal. The point at the edge of each of them specifies the electric field direction at a fixed instant. The open and closed ellipses indicate that the vector of electric field strength at the point (x_0y_0) is rotated clockwise and counterclockwise, respectively.

Figure 9.12a shows the polarization distribution in the second harmonic beam cross section at the nonlinear medium parameters allowing for the presence of one C -line and two L -lines in this cross section. The C -line coincides with the x axis. When passing through it, the rotation angle of the major axis of the polarization ellipse changes stepwise by $\pi/2$. Two L -lines are intersected at the beam center. One of them is vertical, while the other makes an angle of 45° with it. Figure 9.12b was plotted at the parameters of the incident radiation and nonlinear gyrotropic medium that allow for the occurrence of two C -lines and two L -lines in the cross section plane of the signal beam at the double frequency. All these lines pass through the beam center.

Thus, under normal incidence of a uniformly elliptically polarized fundamental Gaussian beam on the surface of an isotropic chiral medium (symmetry group $\infty\infty$), one or two C -lines may occur in the cross section of the reflected second-harmonic beam under any real values of the parameters describing the quadratic optical response of the medium and its surface. The cross section of the second harmonic beam may also contain two, one, or none L -lines. The conditions for the occurrence of polarization singularities in the case of oblique incidence will be considered in the next section.

9.5.3 *Polarization Singularities in Second Harmonic Beam Arising upon Reflection of Elliptically Polarized Gaussian Beam in the Case of Oblique Incidence*

The case of second harmonic generation upon reflection of elliptically polarized beam (9.1) in the case of oblique incidence on the surface of isotropic gyrotropic medium, which is much more complex, is of peculiar interest. In order to obtain the expressions for the electromagnetic field of the reflected wave at double frequency in this case, we have to carry out the following steps [68, 69], generalized the approach described early in Sect. 9.5. In the beginning we solve the equations for the electromagnetic field at the surface of the medium obtained from the conventional boundary conditions (which coincide with the (9.16), when the nonlinear polarization is zero and when we neglect the members of the order of d_0/λ and higher). Thus, we find the electric fields at the surface inside the medium and $\mathbf{E}^{(v)}$ at the surface in vacuum (superposition of the incident and reflected field). After that we substitute these electric fields in (9.17), (9.26) and find the nonlinear polarization of the medium bulk and the nonlinear response of the medium surface. Now we are able to substitute the nonlinear polarization and the nonlinear current density in the modified boundary conditions (9.16), which contain the reflected electric field at second harmonic in the left part of the equations. In order to obtain the electric field strength at second harmonic inside the medium, we solve the wave equation with the nonlinear polarization of the medium bulk in the right part. Finally, we substitute to the boundary conditions (9.16) found above nonlinear polarization, nonlinear current density of the bounded charges, and electric field at second harmonic inside the medium. As a result, we find the reflected wave at double frequency [70]. More details of these steps are given in [68, 69], where this method has been used. Naturally, all mentioned above physical quantities were represented as a superposition of their spatial Fourier harmonics, and all described steps were implemented with their participation.

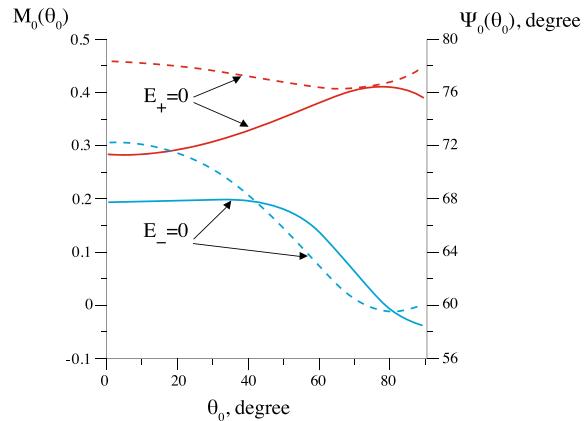
Moreover, we have to assume that the angle of divergence of the light beam at fundamental frequency is small and, therefore, the spatial Fourier harmonics propagating at sufficiently large angle (comparable to 1 rad) relative to the axis of the beam will have negligibly small amplitudes. In this case we can take into account only small values of the transversal components $\mathbf{k}_\perp(\omega)$ of the wave vectors of the spatial Fourier harmonics and to linearize all mentioned in order to obtain the expressions for the electromagnetic field of the reflected wave at double frequency algebraic equations with respect to these components. It is worth noticing that the nonlinear polarization and the nonlinear current density of bounded charges containing the terms of the electric fields look like convolution products of these fields in a Fourier representation with the variable of integration $\mathbf{k}_\perp(\omega)$. Thus, the nonlinear polarization and the nonlinear current density of bounded charges represent the summation of all contributions of all interacting pairs of the Fourier harmonics of the fundamental waves, such that $2\mathbf{k}_\perp(\omega) = \mathbf{k}_\perp(2\omega)$. The last equation imposes a constraint on the variables of integration.

The analysis of expression obtained for the electric field in the second harmonic beam generated from the surface of the isotropic gyrotropic medium by obliquely incident elliptically polarized Gaussian beam (9.1) have shown that the *C*-point appears in the center of the beam if one of the conditions (according to clockwise rotation or counterclockwise rotation of the polarization in the *C*-point of the beam at the double frequency) takes place. These conditions allow one to build dependences $M_0(\theta_0)$ and $\Psi_0(\theta_0)$, where θ_0 is the angle of incidence, for a given set of fixed medium parameters such that, if the polarization of the incident beam follows accordingly to these dependences, the *C*-point remains in the center of the second harmonic beam. Really there are two different pairs of $M_0(\theta_0)$ and $\Psi_0(\theta_0)$ for clockwise and counterclockwise rotation *C*-points. Typical dependences of $M_0(\theta_0)$ (solid line) and $\Psi_0(\theta_0)$ (dashed line) are shown at Fig. 9.13, plotted for the following medium parameters: $n(\omega) = 1.32$, $n(2\omega) = 1.34$, $\omega\gamma_{xxy}^{(2)}(2\omega; \omega, \omega)/b_1 = 1$, $\omega\gamma_{xyx}^{(2)}(2\omega; \omega, \omega)/b_1 = 1.25$, $b_2/b_1 = 0.9$, $b_3/b_1 = 0.2$, and $b_5/b_1 = 2$.

The topological charge of the *C*-point appearing does not change when changing the angle of incidence and a simultaneous change of the polarization state of the incident beam ($M_0(\theta_0)$ and $\Psi_0(\theta_0)$). If the topological charge is positive, the type of the *C*-point can alternate from *lemon* to *monstar* and vice versa. For a fixed value of incident angle, small variations of the polarization state of the fundamental beam do not lead to the disappearance of the *C*-point, but slightly change its position in the second harmonic beam cross-section, moving it from the center.

This can be seen in Fig. 9.14, for which parts (a)–(d) are obtained for $n(\omega) = 1.32$, $n(2\omega) = 1.34$, $\omega\gamma_{xxy}^{(2)}(2\omega; \omega, \omega)/b_1 = 1$, $\omega\gamma_{xyx}^{(2)}(2\omega; \omega, \omega)/b_1 = 1.25$, $b_2/b_1 = 0.9$, $b_3/b_1 = 0.2$, $b_5/b_1 = 0.37$ and slightly different values of $M_0(\theta_0)$ and $\Psi_0(\theta_0)$. These pictures show the light polarization distributions in the cross-section of the reflected beam at the double frequency. Ellipses in different points of the picture have the same shape as the light polarization ellipses in the same points of the beam cross-section. The sum of squared axes of the ellipse is proportional to the

Fig. 9.13 (Adopted from [70]) The dependences of the ellipticity degree M_0 (solid line) and the angle of orientation of the polarization ellipse Ψ_0 (dashed line) of the incident radiation on the angle of incidence θ_0 corresponding to the appearance of the *C*-point (with clockwise ($E_-^{\text{SH}} = 0$) or counterclockwise ($E_+^{\text{SH}} = 0$) rotation) in the center of the second harmonic beam cross-section



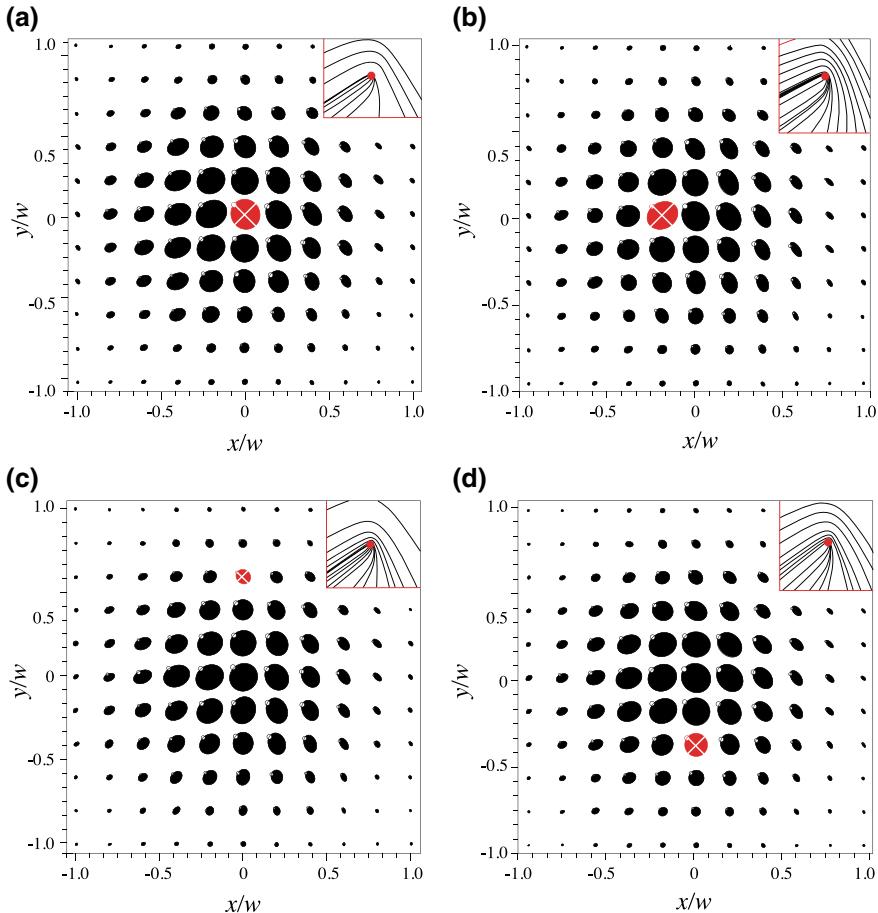


Fig. 9.14 (Adopted from [70]) The polarization distribution in the reflected beam at double frequency at $\theta_0=60^\circ$, $M_0=-0.433$, $\Psi_0=16.5^\circ$ (a); $\theta_0=60^\circ$, $M_0=-0.433$, $\Psi_0=20^\circ$ (b), $\theta_0=60^\circ$, $M_0=-0.5$, $\Psi_0=16.5^\circ$ (c); $\theta_0=58^\circ$, $M_0=-0.433$, $\Psi_0=16.5^\circ$ (d). The singularity point *lemon* is marked with a white cross; the light polarization is circular at this point. Right upper corner insets show the direction of the polarization streamlines (the main axes of the ellipses are tangent to these lines at each point) in the vicinity of the C-point

intensity of light, which is normalized at maximum intensity in each picture. The dot at the edge of each of the ellipses indicates the position of the end of the vector of electric field strength at fixed timing. For all parts of Fig. 9.14, the sense of rotation of the electric field vector remains the same (counterclockwise) along the cross-section of the beam.

Figure 9.14a is plotted for $\theta_0=60^\circ$, $M_0(60^\circ)=-0.433$ and $\Psi_0(60^\circ)=16.5^\circ$, therefore a *lemon* C-point is located in the center of the beam. It is marked with a cross at the gray ellipse. The direction of the polarization streamlines (the main axes of the ellipses are tangent to these lines in each point) in the vicinity of the C-point

is shown in the right upper corner inset. Parts (b)–(d) of Fig. 9.14 are plotted for values of θ_0 , M_0 and Ψ_0 slightly varying from those corresponding to the appearance of the C -point in the center of the beam in Fig. 9.14a: for (b) we changed Ψ_0 to 20° retaining $\theta_0 = 60^\circ$ and $M_0 = -0.433$; for (c) we changed M_0 to -0.5 ; for (d) we changed θ_0 to 58° . In all these three cases the C -point moves out of the center of the beam. It is worth noting that separate changes of M_0 , Ψ_0 and θ_0 move the C -point in different directions. The insets in the right upper corner show polarization streamlines in the vicinity of the C -point in each case.

Let pay special attention to the case when the angle of incidence is small. Figure 9.15 shows the inhomogeneous polarization distribution at the second harmonic beam for $n(\omega) = 1.32$, $n(2\omega) = 1.34$, $\omega\gamma_{xxy}^{(2)}(2\omega; \omega, \omega)/b_1 = 2$, $\omega\gamma_{xxy}^{(2)}(2\omega; \omega, \omega)/b_1 = 2.5$, $b_2/b_1 = 0.9$, $b_3/b_1 = 0.2$, and $b_5/b_1 = 2$. For all parts of this figure the sense of rotation of the vector of electric field strength remains clockwise along the cross-section of the beam, and this is indicated by unfilled (white) ellipses in opposite to Fig. 9.14, where all the ellipses are shaded. If $\theta_0 = 60^\circ$ (Fig. 9.15a), then reflected second harmonic beam attains additional symmetry (the polarization state of light in this beam depends only on the polar angle coordinate and not on the polar radius (see Sect. 9.5.2)). The ellipticity degree and the angle of orientation of the polarization ellipse remain unchanged along the arbitrary straight line in the cross-section of the signal beam crossing the center of the beam (the intensity is zero in the center). There no point singularities can appear in the cross-section of the signal beam, only L -lines and C -lines. In Fig. 9.15a, plotted for $M_0 = 0.195$ and $\Psi_0 = 72.34^\circ$, the C -line highlighted by gray color appears as a horizontal line. For any other value of Ψ_0 the polarization distribution and the C -line will be rotated by an angle $\Psi_0 = 72.34^\circ$ around the center of the beam (z -axis) without any change in its structure due to the symmetry of the medium.

Now let us incline the incident beam at a small angle θ_0 from the vertical position in such a way that the beam axis, the C -line (in the normally reflected signal beam) and the perpendicular to the surface remain in the same plane (see Fig. 9.16). After inclining the incident beam, one can see that C -point appears close to the center of the reflected second harmonic beam (Fig. 9.16b). This can be explained by the fact that for this small value of θ_0 “current” values of $M_0 = M_{0-}$ and $\Psi_0 = -\varphi_{C-}$ (see Sect. 9.5.2) are still close although not equal to the $M_0(\theta_0)$ and $\Psi_0(\theta_0)$.

When increasing the angle of incidence (Fig. 9.16c) the C -point moves away from the center along the direction which coincides with the direction of the C -line in the cross-section of the normally reflected beam at the double frequency. This can be qualitatively explained for small θ_0 . A small variation of the angle of incidence near to normal incidence ($\theta_0 = 0$) is really equivalent to a variation of the transverse component of the wave vector in the Fourier representation of the light beam. Thus, we efficiently change the direction of the wave vector of the spatial Fourier component responsible for the appearance of the singularity in the beam. As a result the C -point moves away from the center, in a direction which is opposite to the inclination direction, i.e., in the positive direction of the x -axis.

When inclining the incident beam at any other direction which does not coincide with the C -line (by an appropriate choice of θ_0 determining the position of the C -line, for example, perpendicular to the x -axis, as in Fig. 9.16d, $\theta_0 = 1^\circ$, $\Psi_0 = -17.66^\circ$), the value of Ψ_0 is too far from $\Psi_0(\theta_0=0)$ and the C -point does not appear in the reflected signal beam.

9.5.4 Singular Polarization Patterns of the Beam at Double Frequency Generated by Singularly Polarized Fundamental Beam

In this section we consider the formation of the lines of circular polarization in the second harmonic beam's light field reflected from the surface of an isotropic gyrotropic medium in the case when normally incident fundamental beam contains a polarization singularity [71]. Let the incident beam at frequency ω consists of the coaxial Gaussian mode (9.1) with left-hand circular polarization and two Laguerre-Gaussian modes of the first order:

$$\mathbf{E}_l = \frac{E_{0L}}{w\beta^2(z)} \left[\mathbf{e}_+ + ik^{-1} \mathbf{e}_z (\mathbf{e}_+ \cdot \vec{\nabla}_\perp) \right] [p(x+iy) + q(x-iy)] \times \exp\left(-\frac{x^2+y^2}{w^2\beta(z)} - i\omega t + ik(z-l_0)\right) \quad (9.50)$$

with right-hand circular polarization. Here E_{0L} is the amplitude, $\beta(z) = 1 + 2i(z - l_0)/(kw^2)$, $p = \cos(\theta/2)$, $q = \exp(i\eta) \cos(\theta/2)$, angle θ takes the values from 0 to π and governs the amplitude ratio of two Laguerre-Gaussian modes in the incident beam. The parameter η depends on the difference of the phases of these two modes as well as on the phase shift of the Gaussian mode. It determines only the topological type of the left-hand polarization singularity (*lemon* or *monstar*) at the Oz axis in case of positive value of its topological index $I_C = 0.5 \text{sgn}\{\cos \theta\}$. For the simplicity we assume the modes to be focused at the surface of the nonlinear medium and to have the same waist sizes w . The examples of the transversal polarization distribution in the waist of the fundamental beam at various θ and η are given in Fig. 9.17. Empty ellipses correspond to the right-hand polarized radiation and filled ones correspond to the left-hand polarized radiation.

Carrying out the same calculations as in the previous sections one can obtain analytical expressions for the circularly polarized components of the electric field strength of the second harmonic beam:

$$E_+^{SH} = A(\rho, z) \{ \rho \exp(-i\varphi) D_{2+} + h[(D_0 - D_{2+})(\rho^2 \beta^{-1}(z) K(\varphi) \exp(-i\varphi) - 0.5 \cos(\theta/2)) + 0.5 D_{1+} \beta(z) \cos(\theta/2)] \}, \quad (9.51)$$

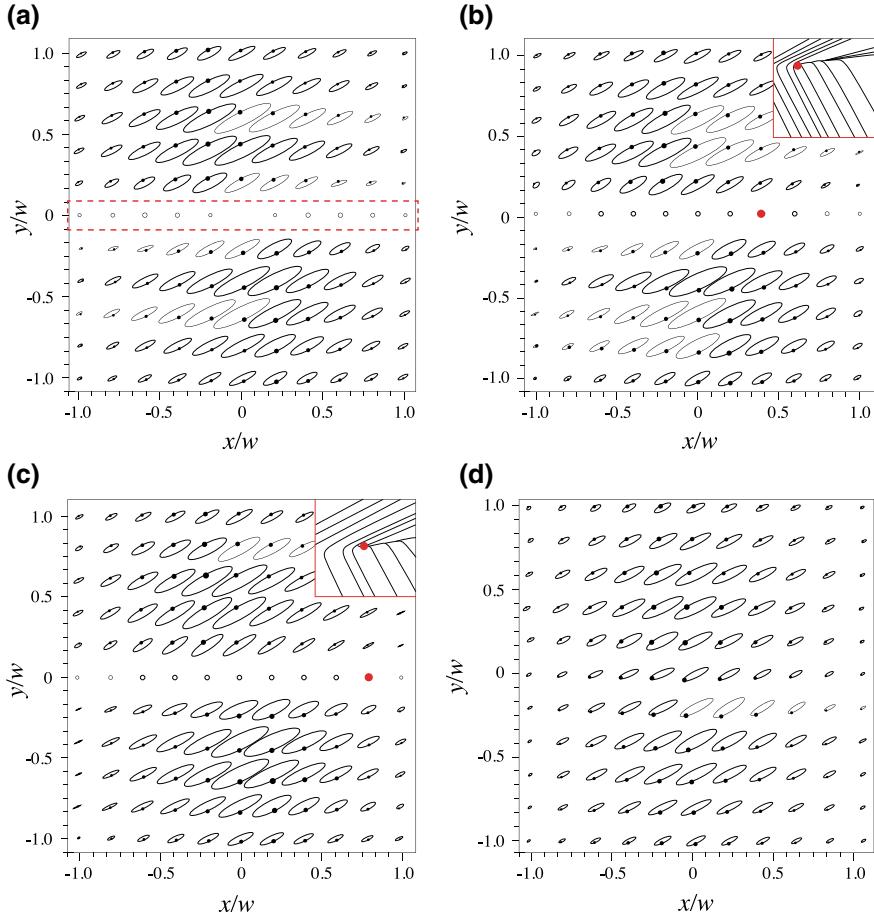


Fig. 9.15 (Adopted from [70]) The polarization distribution in the reflected beam at double frequency: **a** normal incidence, the C-line is highlighted by a rectangle shown by broken line; **b** oblique incidence $\theta_0 = 1^\circ$, such that the axis of the incident beam lies in the same plane as the perpendicular to the surface and the C-line shown in (a); the C-line is transformed into the C-point; **c** further increase of the angle of incidence ($\theta_0 = 2^\circ$) shifts the appearing C-point away from the center of the beam; **d** oblique incidence $\theta_0 = 1^\circ$ when the beam axis is inclined perpendicularly to the C-line direction (shown in (a)). No singularity appears in the central bright region of the beam

$$\begin{aligned}
 E_-^{SH} = & hA(\rho, z) \left\{ h\rho\beta^{-1}(z)D_{2-} - [\rho^2\beta^{-1}(z)K^2(\varphi)\exp(i\varphi) - L(\varphi, z)] + (D_0 - D_{2-}) \right. \\
 & \times [\rho^2\beta^{-1}(z)K(\varphi)\exp(-i\varphi) - 0.5\exp(i\eta)\sin(\theta/2)] - 0.5D_{1-}\beta(z)\exp(i\eta)\sin(\theta/2)] \left. \right\}. \tag{9.52}
 \end{aligned}$$

In these formulae $\rho = (x^2 + y^2)^{1/2}/w$, the functions A , K , L have the following form:

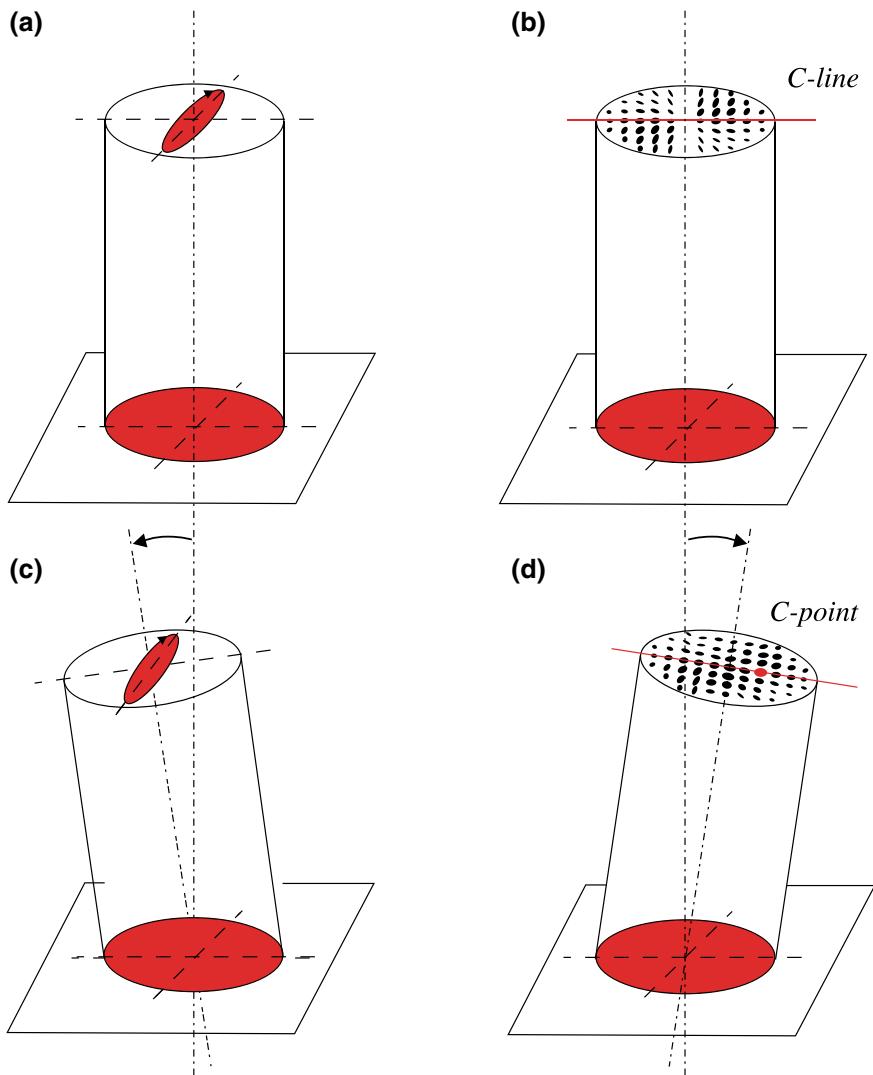


Fig. 9.16 (Adopted from [70]) The scheme of proper inclination of the incident fundamental beam and the corresponding inclination of the reflected second harmonic beam, when the *C*-line in the normally reflected second harmonic beam transforms into the *C*-point: **a** normally incident fundamental beam and **b** normally reflected signal beam with the *C*-line in the cross-section. **c** Oblique incident fundamental beam (the axis of the beam, the *C*-line in the normally reflected beam and the perpendicular to the surface remain in the same plane) and **d** the reflected signal beam with the *C*-point, which is the result of the transformation of the *C*-line

$$A(\rho, z) = \frac{32i\sqrt{2}\pi E_{0L}^2}{\beta^2(z)\omega w[1+n(\omega)]^2[1+n(2\omega)]} \exp\left(-\frac{2\rho^2}{w^2\beta(z)}\right) \quad (9.53)$$

$$K(\varphi) = \exp(-i\varphi)\cos(\theta/2) + \exp[i(\varphi+\eta)]\sin(\theta/2), \quad (9.54)$$

$$L(\varphi, z) = 0.5[(1 + \cos\theta)\exp(-i\varphi) + \exp[i(\varphi+\eta)](2 - \beta(z))\sin\theta], \quad (9.55)$$

where $D_0 = n(2\omega)b_1 + i\omega\gamma_{xyy}^{(2)}(2\omega; \omega, \omega)/n(2\omega)$, $D_{1\pm} = D_{2\pm} + i\omega\gamma_{xyy}^{(2)}(2\omega; \omega, \omega)/[n(\omega) + n(2\omega)]$, $D_{2\pm} = n(\omega)(b_3 \pm ib_5)$, $h = E_{0L}/E_0$.

The positions and amount of the *C*-points in the cross-section of the beam at the double frequency vary together with the coordinate z . In certain cases a pairwise annihilation occurs of two *C*-points with the same direction of the electric field rotation but opposite topological indices above the surface of the nonlinear medium. The inverse process is also possible: two *C*-points of opposite topological indices can be nucleated. To find the positions of *C*-points and the shapes of *C*-lines one has to solve the equations $E_\pm^{\text{SH}} = 0$. To obtain analytical solutions of these equations is possible only in some important particular cases which will be considered below.

If the linear and nonlinear surface response is negligibly small ($\hat{k}^{(1)} = \hat{k}^{(2)} = 0$), the changes in η cause only the rotation of the signal beam around z axis as a whole but do not change its polarization distribution and it gives ability to assume $\eta = 0$. If $\theta < \pi/2$ (positive index of the *C*-point in the fundamental beam) in certain transversal planes z of the second harmonic beam there are two pairs of left-hand *C*-points. The points of each pair are located symmetrically with respect to the beam's axis and the singularities of the one pair have positive topological index (1/2), while the singularities of the other pair have negative index (-1/2). Parameter $\gamma_r = [1 + n(2\omega)/n(\omega)][\gamma_{xyy}^{(2)}(2\omega; \omega, \omega)/\gamma_{xyy}^{(2)}(2\omega; \omega, \omega)]$ determines the intervals of the propagation coordinate z , in which these singularities exist. At the boundaries of these intervals the pairwise annihilation or nucleation of the singularities takes place.

If the *C*-point's index of the incident beam is positive ($\theta < \pi/2$) whatever the values of z , θ and γ_r are, there is a pair of right-hand *C*-points in each cross section of the signal beam. They lie symmetrically with respect to the beam's axis and have the indices 1/2. When $\theta = 0$, these two *C*-points merge at the beam's axis into the single one with index 1. We notice that both *C*-points have the same topological index and direction of the electric field's rotation as the *C*-point in the reflected beam at fundamental frequency. If $\theta > \pi/2$ then the beam at doubled frequency has the structure analogous to the described one for the case $\theta < \pi/2$. The beam contains two left-hand singularities, the index of each of which is still 1/2. If $\theta = \pi$ then these two *C*-points merge into one with index 1. Four additional right-hand singularities arise if the parameters z , θ and γ_r are linked by certain relationships. The topological indices of these four *C*-points are in pairs opposite.

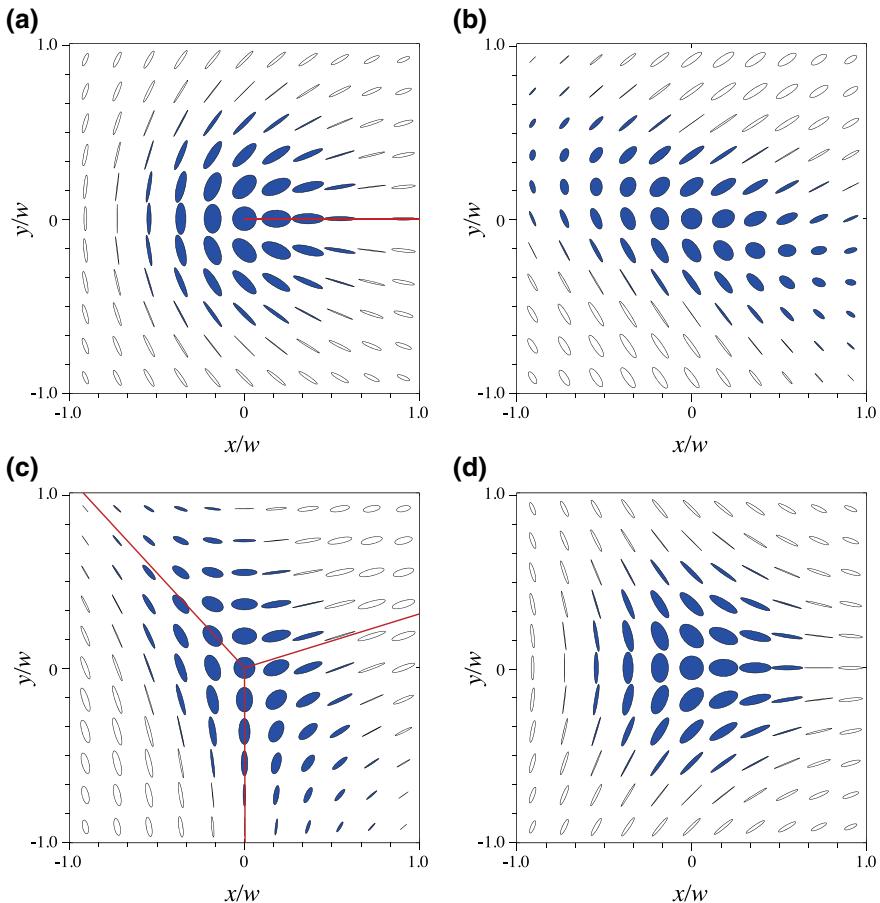


Fig. 9.17 (Adopted from [71]) The examples of the transversal polarization distribution in the waist of the fundamental beam near the C-point at $\theta=0$, $\eta=0$ (a), $\theta=0.4\pi$, $\eta=3\pi/4$ (b), $\theta=0.7\pi$, $\eta=\pi/3$ (c), $\theta=\pi$, $\eta=0$ (d). The lines designate the polarization lines starting from the C-point which lies at the origin

The doubling of the sum topological index, observed at any θ , is analogous to the doubling of the vortex's topological charge (see [72]). However, it is worth mentioning that when the fundamental beam's C-point has negative index ($\theta>\pi/2$) the sum topological index is not only doubled but also changes its sign. When $\theta<\pi/2$ the direction of the electric field's rotation in these “uncoupled” singularities coincide with this of the reflected beam at fundamental frequency, but when $\theta>\pi/2$ the directions of rotation are opposite. Figure 9.18 illustrates the transversal polarization distributions in the second harmonic beam for some of the described cases. The C-points with positive (negative) topological indices are marked by fill (empty) circles in this figure.

In general case of nonlinear media, the surface response of which is as much significant as the nonlocal bulk response ($\hat{\kappa}^{(1)} \neq 0$, $\hat{\kappa}^{(2)} \neq 0$), the problem becomes far more complicated. In this case we search the analytical expressions and the dynamics of the generated singularities for two specific configurations of the incident beam ($\theta=0$ and $\theta=\pi$). In the first case from (9.52) one can find that three right-hand C -points in each cross-section of the second harmonic beam are. One point always lies at its axis ($\rho_{1-}(z)=0$) and has the index of 1/2, while the positions of two other are given by the formulae:

$$\rho_{2,3-}(z) = \left[1 + 0.5\lambda|\beta(z)|^2 \pm \sqrt{1 + (\lambda - 1)|\beta(z)|^2 + 0.25\lambda^2|\beta(z)|^4} \right]^{1/2}, \quad (9.56)$$

$$\varphi_{2,3-}(z) = \arg \left\{ (D_0 - D_{2-})^* D_{2-} \right\} - \arg \left\{ \rho_{2,3-}^2(z)/\beta(z) - 1 \right\}. \quad (9.57)$$

Here $\lambda = |(D_{2-} - D_0)/(hD_{2-})|^2 > 0$. The topological indices of these two C -points are $-1/2$ and $1/2$. If $h < \sqrt{2}|D_{2-} - D_0|/|D_{2-}|$, then two values $\rho_{2,3-}$ are real at any z and the beam at the double frequency contains three C -lines, along which the polarization is right-hand. If $h > \sqrt{2}|D_{2-} - D_0|/|D_{2-}|$, there is an area of space $z \in [z_a, z_n]$, in which the solutions $\rho_{2,3-}$ are complex. At the boundaries of this interval, the pairwise annihilation (index “ a ”) or nucleation (index “ n ”) of the singularities takes place. The values $z_{a,n} = (kw^2/2\lambda)(2 - 2\lambda - \lambda^2 \mp [1 - 2\lambda]^{1/2})$. Inside this interval only trivial C -line $\rho_{1-}(z)=0$ exists. In the planes $z=z_{a,n}$ at the points with coordinates $\rho=\rho_{2-}=\rho_{3-}$ and $\varphi=\varphi_{2-}=\varphi_{3-}$ the pairwise annihilation ($z=z_a$) or nucleation ($z=z_n$) of these two C -points takes place. The sum topological index of the right-hand C -points is $1/2$, unlike the case of purely bulk medium’s response. However, as D_{2-} tends to zero the singularity of negative index (upper sign “+” in (9.56) and (9.57)) moves away from the axis of the second harmonic beam.

The solutions of $E_+^{\text{SH}}=0$, which determine the positions of the left-hand singularities, are linked by the following relationships when $\theta=0$:

$$\begin{aligned} \rho_+(z) = & (D_{2+} \exp(-i\varphi_+) \pm \{D_{2+}^2 \exp(-2i\varphi_+) - 2h^2(D_0 - D_{2+})(D_1 + \beta(z) + D_{2+} - D_0) \\ & \times \beta^{-1}(z)\}^{1/2}) [2h(D_{2+} - D_0)\beta^{-1}(z)]^{-1}. \end{aligned} \quad (9.58)$$

Two C -points, corresponding to real $\rho_+(z)$, have opposite topological indices. Numerical investigations of the relations (9.58) show there can exist the intervals of z coordinate, in which ρ_+ cannot be real at any φ_+ and singularities do not exist in such cross-sections. Figure 9.19 shows a typical pattern of C -lines and the transversal polarization distributions in the second harmonic beam. C -lines,

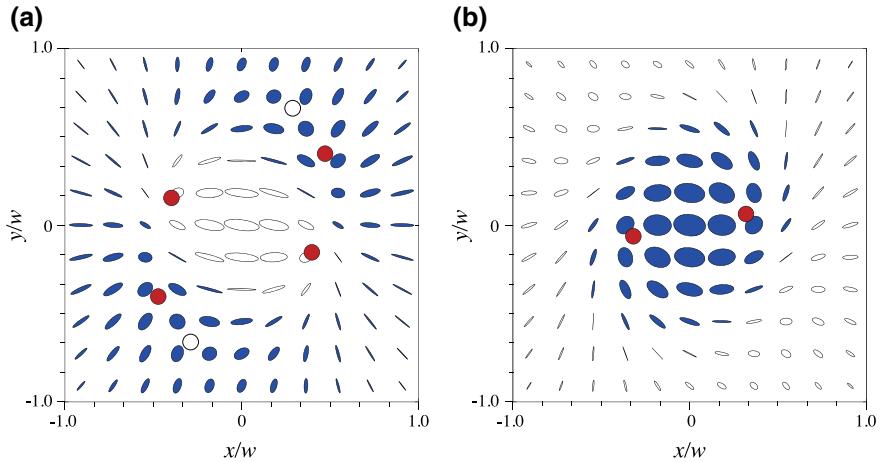


Fig. 9.18 (Adopted from [71]) Transversal polarization distributions in the signal beam at distance $kw^2/4$ above the surface of the nonlinear medium with purely bulk response at $\theta=\pi/5$ (a), $\theta=4\pi/5$ (b) and $h=1$, $\eta=0$. The parameters of the nonlinear medium are $n(\omega)=1.33$, $n(2\omega)=1.35$, $b_1=b_3=b_5=0$, $\gamma_{xyxy}^{(2)}/\gamma_{xxyy}^{(2)}=0.5$

corresponding to the singularities of positive (negative) index, are designated by full (dashed) lines, white stars mark the places of the singularities' pairwise nucleation and the black ones mark the places of pairwise annihilation.

When $\theta=\pi$ the polarization distribution in the cross-section of the incident beam has the rotation symmetry axis of the third order (see Fig. 9.17d). The analysis of (9.51) when $\theta=\pi$ shows that left-hand C-points lie at the vertices of an equilateral triangle in each cross-section of the signal beam, the center of the triangle lying at the origin. Each of the corresponding C-points has the topological index of 1/2. There is also the C-point at the axis of the second harmonic beam, the index of which is $-1/2$. The sum topological index of the left-hand singularities is 1, like in the case of purely bulk response. When D_{2+} tends to zero three C-points with positive indices move towards the beam's axis, and merge with the central singularity into C-point with index "1" when $D_{2+}=0$.

The expressions for the coordinates of the right-hand C-points, which are obtained from the equation $E_-^{\text{SH}}=0$ have very cumbersome form. The analysis of them allows one to conclude that there are one or three different values of ρ_- , at which the equation $E_-^{\text{SH}}=0$ becomes an identity at three values of φ_- , which differ from each other in $2\pi/3$. The sum topological index of all right-hand singularities is $-3/2$ in any case. Figure 9.20 illustrates one of three symmetrically located C-lines structure with the central C-line and the transversal polarization distributions of the signal beam when $\theta=\pi$.

9.6 Formation and Interaction of Polarization Singularities During Light Self-Action and Interaction of Elliptically Polarized Beams in Isotropic Medium with Spatial Dispersion of Cubic Nonlinearity

9.6.1 Ring-Shaped Electric Field Structures Formation During Self-Focusing of Elliptically Polarized Gaussian Beam

The self-focusing of linearly polarized light is most frequently investigated in the aberrationless (paraxial) approximation, allowing for the constancy of the form of the function describing the shape of a Gaussian beam during the course of its propagation in a nonlinear medium [73, 74]. The application of this approximation yields a good agreement with experimental results and numerical computations for the initial stage of collapse of the beam near its axis [75]. The use of the generalization of the approximation [76], allowing the use of the method of moments [77] as well, made it possible to ascertain that the nonlocality of the nonlinear optical response of matter causes substantial changes in the threshold conditions of self-focusing and other propagation regimes of elliptically polarized beams of a Gaussian profile. A comparison of the spatial-coordinate dependences of the beam intensity, degree of ellipticity, and rotation angle of the major axis of the polarization ellipse, with the results of numerical investigation of the system of connected parabolic nonlinear diffraction equations describing the self-action of light in such a medium, shows them to be in good qualitative agreement only in the vicinity of the beam axis. The slight quantitative discrepancy [76] between the positions of the boundaries of the regions in the space of the incident radiation parameters, within which similar propagation regimes occur, which was found numerically and with the

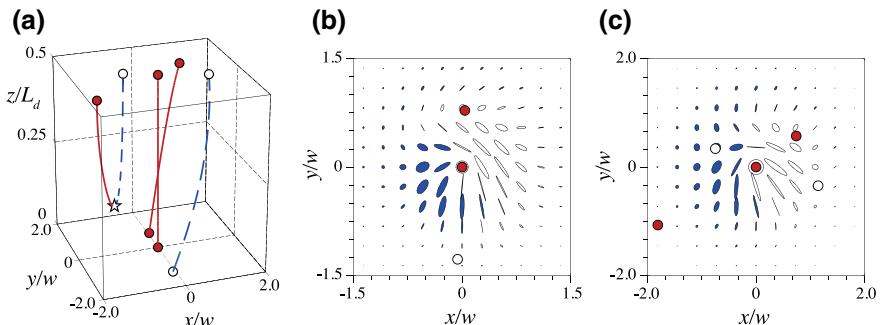


Fig. 9.19 (Adopted from [71]) C-lines (a) and the transversal polarization distributions of the signal beam at the surface of the medium (b) and at the distance $kw^2/4$ above it (c). The incident beam's parameters are $\theta = 0$, $h = 1$, $\eta = 0$, the parameters of the nonlinear medium are $n(\omega) = 1.33$, $n(2\omega) = 1.35$, $b_3/b_1 = 0.75$, $b_5/b_1 = 0.3$, $\omega\gamma_{xyy}^{(2)}/b_1 = 0.5$, $\omega\gamma_{xxy}^{(2)}/b_1 = 0.25$

aid of the approximation used in [76, 77]. This is explained by the difference between the beam collapse criteria used, with the transverse spatial intensity and polarization distributions within the radially symmetric propagating beam being strongly different from the respective Gaussian and uniform distributions at the boundary of the medium.

A consistent investigation of the changes in the distributions of intensity, degree of ellipticity, and rotation angle of the major axis of the polarization ellipse that occur in the cross-sectional plane of a light beam suffering self-action in an isotropic medium with spatial nonlocality of optical response has revealed a most interesting effect — the formation of circularly polarized ring-shaped light field structures. The description their formation dynamics is the subject of this section.

The propagation of an elliptically polarized light beam in an isotropic media with spatial dispersion of cubic nonlinearity is described by the following system of parabolic equations [11, 76] for the slowly varying complex circularly polarized wave amplitudes $A_{\pm}(r, z)$:

$$\frac{\partial A_{\pm}}{\partial z} + \frac{i}{2k} \Delta_{\perp} A_{\pm} = \pm i\rho_0 A_{\pm} - i \left\{ (\sigma_1/2 \mp \rho_1) |A_{\pm}|^2 + (\sigma_1/2 + \sigma_2) |A_{\mp}|^2 \right\} A_{\pm}. \quad (9.59)$$

Here $\sigma_{1,2}$ are proportional to the components of the local nonlinear susceptibility tensor $\hat{\chi}^{(3)}(\omega; \omega, \omega, -\omega)$: $\sigma_1 = 4\pi\omega^2 \chi_{xxyy}^{(3)}/kc^2$, $\sigma_2 = 2\pi\omega^2 \chi_{xyyx}^{(3)}/kc^2$. The pseudoscalar constant $\rho_0 = 2\pi\omega^2 \gamma_0/c^2$, where $\gamma_{ijk}^{(1)} = \gamma_0 e_{ijk}$. The nonzero component $\gamma_{xxyyz}^{(3)}$ of the nonlocal cubic susceptibility tensor $\hat{\gamma}^{(3)}(\omega; \omega, \omega, -\omega_1)$ is proportional to $\rho_1 = 2\pi\omega^2 \gamma_{xxyyz}^{(3)}/c^2$. We will assume that a light beam of peak intensity I_0 and the degree of ellipticity M_0 incident upon the boundary of the medium has a Gaussian form:

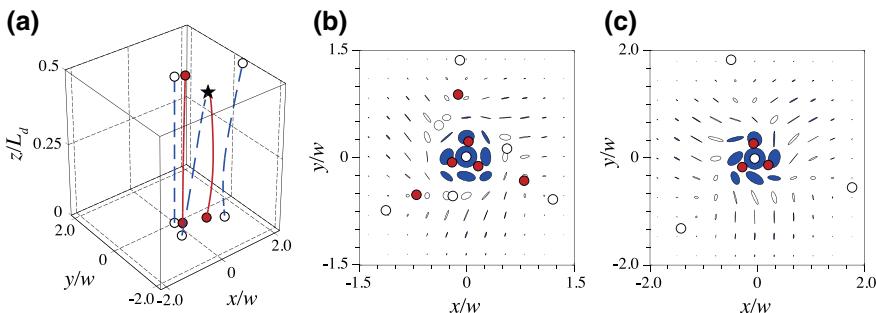


Fig. 9.20 (Adopted from [71]) C-lines (a) and the transversal polarization distributions of the signal beam at the surface of the medium (b) and at the distance $0.4kw^2$ above it (c). The incident beam's parameters are $\theta = \pi$, $h = 1.6$, $\eta = 0$, the parameters of the nonlinear medium are $n(\omega) = 1.33$, $n(2\omega) = 1.35$, $b_3/b_1 = 0.15$, $b_5/b_1 = 0.3$, $\omega r_{xxyy}^{(2)}/b_1 = 0.5$, $\omega \gamma_{xxyy}^{(2)}/b_1 = 0.25$

$$A_{\pm}(r, z=0) = \left[I_0 (1 \pm M_0)^{1/2} \right] \exp(-r^2/r_0^2). \quad (9.60)$$

At a distance z from the boundary the characteristics of the propagating beam depend on the dimensionless power $P = \sigma_1 L_d I_0$ and M_0 and also on the following parameters characteristic of the medium: $\rho_0 L_d$, σ_2/σ_1 , and ρ_1/σ_1 . The formula for P includes the diffraction length: $L_d = kr_0^2/2$. We will further assume that $\sigma_1 > 0$.

The formation of a ring-shaped light field structure in a nonlinear medium is due to the difference in the r dependence between the expressions in the curly brackets in (9.59). The first terms are always positive (the spatial dispersion of the media of interest is considered to be small) and make a “focusing” contribution to the propagation of the circularly polarized light field components A_{\pm} . The second terms between the curly brackets in formula (9.59) are negative if $\sigma_2/\sigma_1 < -0.5$. In that case, they make a “defocusing” contribution to the propagation of A_{\pm} . The opposite signs of these terms give rise to additional extrema in the right-hand members of formula (9.59). The consequence is the formation of ring-shaped or even more complex radially symmetric light field structures in the cross-sectional plane of the light beam having a uniformly distributed polarization mode and normally distributed intensity (9.60) at $z=0$.

Figure 9.21 presents some examples of such structures developing at $\sigma_2/\sigma_1 < -0.5$. One can easily see that the propagating beam is inhomogeneously polarized: one can find linearly polarized, elliptically polarized, and circularly polarized beam regions, with the rotation of the electric field changing sense with increasing r . The amount of the emerging radially symmetric ring-shaped structures uniting the points of the same sense of rotation of the electric-field vector most strongly depends on P . Three such ring structures can readily be seen in Fig. 9.21a, and seven, in Fig. 9.21b. At $|\sigma_2/\sigma_1| \leq 2$ the transverse structure of the light field, similar to that shown in Fig. 9.21a, is formed in accordance with the following scenario. The “weaker” (in intensity) circularly polarized component of the propagating wave having a Gaussian form at $z=0$ starts decreasing at the point $r=0$ as z increases. At a certain z/L_d a maximum is formed at a sufficiently short distance from the point $r=0$. As z further increases, the magnitude of this maximum also starts to smoothly diminish, and the point at which it is reached moves away from the beam axis. In some cases, another intensity maximum can form at a short enough distance from the point r_1 that will behave similarly with increasing z . As a result, the intensity distribution of the weaker, circularly polarized electric field component assumes a ring-like form. The “stronger,” at $z=0$, circularly polarized component retains, in the course of propagation, its Gaussian form with a maximum at $r=0$.

Recall that the propagation of a light beam in a nonlinear medium occurs without any change in its degree of ellipticity [76], provided that $M_0 = -\rho_1/\sigma_2$. At $1 > M_0 > -\rho_1/\sigma_2$ the “stronger” component is A_+ , and, starting with some z value,

$M(r=0, z)$ becomes equal to unity. If the inequality $-1 < M_0 < -\rho_1/\sigma_2$ holds true, it is A_- that the “stronger” component. As z increases, the degree of ellipticity at the beam axis tends to -1 . The smaller is the difference between M_0 and ± 1 , the sooner the distribution becomes ring-like. On completion of the formation of the ring-shaped structure, the maximum intensity of the “weaker” component amounts to less than 5% of the peak intensity of the “stronger” component. The bright (with respect to the e^{-2} level) central part of the beam recorded at the exit from the medium is almost circularly polarized: $|M(r=0, z)|$ differs from unity by less than one-tenth of a percent. The light field of the “weaker” component is in this case distributed over a considerably greater area.

The scenario for the formation of ring-shaped light field structures at high $|\sigma_2/\sigma_1|$ values is more complicated. At small z_1 the r dependence of the “weaker” component features several extrema. At the point $r=0$ both a local maximum and a local minimum of its intensity can now form. The “stronger” component here does not necessarily reach its absolute maximum at the point $r=0$, with its collapse being preceded by a slight defocusing. The transverse spatial distributions of the normalized intensity I/I_0 (Fig. 9.22a) and degree of ellipticity (Fig. 9.22b) of the light field prove to be rather complex. In the general case, $|M(r=0, z/L_d)|$ varies in a nonmonotonic fashion with increasing z/L_d and does not always tend to unity.

The possibility of C -points formation in the elliptically polarized laser beam, focused into nonlinear medium are of special interest. The singularities did not appear during light propagation in considered above specific range of the nonlinear medium’s parameters. It was numerically and analytically shown [78] that polarization singularities can emerge when a homogeneously elliptically polarized light

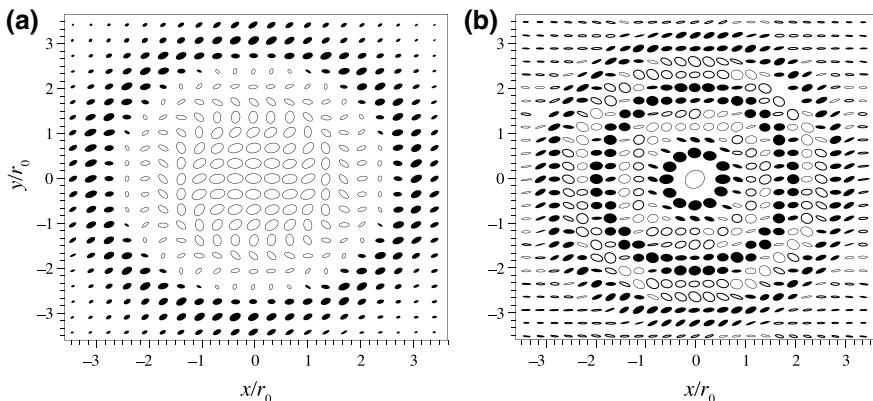


Fig. 9.21 (Adopted from [11]) Radiation intensity and polarization distributions over the cross-section of the self-focusing beam at $P=15$, $\sigma_2/\sigma_1 = -5$, $\rho_1/\sigma_1 = 0$, $M_0 = 0.2$, $z/L_d = 0.4$ (a) and $P=50$, $\sigma_2/\sigma_1 = -4$, $\rho_1/\sigma_1 = 0.2$, $M_0 = 0.05$, $z/L_d = 0.5$ (b)

beam undergoes self-focusing in an isotropic third-order Kerr medium without frequency and spatial dispersion. In the case of axial symmetry of incident beam's intensity profile the generation of *C*-lines takes place in separate planes, perpendicular to the beam's axis. The lines have the shape of circumference and it was shown a single *C*-line generation. The closer the polarization state of initial beam is to the circular, the closer to the medium's border the line is. If the axial symmetry of the incident beam's intensity profile is broken, then there is a domain of propagation coordinate, in which two pairs of *C*-points exist and their total topological charge is zero. The growth of asymmetry parameter of the beam leads to the expansion of this domain. The presence of small amplitude noise in the incident beam does not principally influence the dynamics of formation and propagation of the *C*-points. The latter gives one a chance to detect these singularities in a real experiment.

Numerical simulations of initially homogenously polarized Gaussian light beam self-focusing in isotropic phase of nematic liquid crystal, the temperature of which is close to the temperature of nematic-isotropic phase transition, gives more complicated picture [79]. *C*-lines of both polarization handednesses are nucleated near these transversal cross-sections of the beam, in which local extrema of peak intensities of corresponding circularly polarized components of the beam are attained. These lines are closed loops, surrounding the beam axis. They appear in wide range of beam and nonlinear medium parameters.

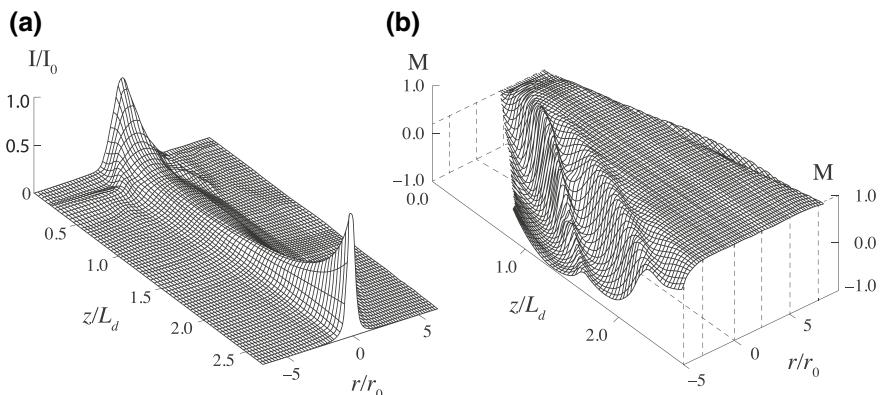


Fig. 9.22 (Adopted from [11]) Variation of the transverse spatial distributions of normalized intensity (a) and degree of ellipticity (b) with increasing propagation coordinate at $P=8$, $\sigma_2/\sigma_1=-5$, $\rho_1/\sigma_1=0$, $M_0=0.2$

9.6.2 Interaction of Two Monochromatic Beams with Polarization Singularities in Case of Their Collinear Propagation

In this section, we consider [80] pairwise creation and annihilation of the *C*-points with opposite topological charges in the process of interaction of two collinear, specific-kind coherent monochromatic beams at frequency ω with polarization singularities in the nonlinear isotropic chiral medium (symmetry group $\infty\infty$). Let us assume two beams falling along the z -axis onto the medium surface $z=0$ normally. The beams have the following structure: the right-handed circularly polarized component has Gaussian profile, while the left-handed circularly polarized one has Laguerre–Gaussian profile. Their centers are located at the x -axis symmetrically with respect to the point $x=0$ at distance d from each other. The slowly varying envelopes $A_{+1,2}$ of right-handed circularly polarized components of these two beams at $z=0$ are given by the following expressions:

$$A_{+1}(x, y) = E_0 \exp\left(-\frac{(x+d/2)^2 + y^2}{w^2}\right), \quad (9.61)$$

$$A_{+2}(x, y) = E_0 \exp(i\theta) \exp\left(-\frac{(x-d/2)^2 + y^2}{w^2}\right), \quad (9.62)$$

and the slowly varying envelopes $A_{-1,2}$ of left-handed circularly polarized components are given as:

$$A_{-1}(x, y) = hE_0 \left(\frac{x+d/2 \mp iy}{w}\right) \exp\left(-\frac{(x+d/2)^2 + y^2}{w^2}\right), \quad (9.63)$$

$$A_{-2}(x, y) = hE_0 \exp(i\theta) \left(\frac{x-d/2 \mp iy}{w}\right) \exp\left(-\frac{(x-d/2)^2 + y^2}{w^2}\right). \quad (9.64)$$

Here E_0 is the amplitude, $\exp(i\theta)$ determines the phase shift between the beams, h is a real value coefficient. When $d \gg w$ both beams have radial symmetric intensity distributions and *C*-points in their centers $(-d/2, 0, 0)$ and $(d/2, 0, 0)$. The radiation is right-handed circularly polarized in these *C*-points. If there is a minus in the expressions (9.63) and (9.64) before the imaginary unit, then the topological index of the *C*-point is $1/2$, and if there is a plus in these equations, then the topological index is $-1/2$. We limit ourselves by the study of topological charge evolution, nucleation and annihilation of the *C*-points with opposite topological charges.

The system of (9.59) is symmetric relative to the simultaneous change of A_{\pm} to A_{\mp} and $\rho_{0,1}$ to $-\rho_{0,1}$. In this case the indices of the *C*-points of the incident beams

change their sign to the opposite in each point of space. Moreover, that simultaneous change of the topological indices of the C -points in the incident beams (which can be achieved by the change of ‘ \mp ’ to ‘ \pm ’ in (9.63) and (9.64)) is equivalent to the reversion of the y -axis, i.e., the spatial distribution of the electromagnetic field and the C -lines are mirror reflected from the plane $y=0$. The effect of linear optical activity ($\rho_0 \neq 0$) does not depend on the transversal coordinates and the topological charge of the C -points and their positions do not change. Therefore, we do not consider the effects connected with linear optical activity, assuming $\rho_0 = 0$. It does not prevent us from accounting for the linear rotation at any stage of light propagation by rotating each ellipse in each point of the beam cross-section by the same angle. The C -points with the opposite handedness of the polarization rotation do not interact in a linear medium. Thus, we can limit ourselves by the consideration of incident beams with right-handed (or left-handed) circular polarization in both C -points with equal or opposite topological indices. In these two cases the equation system (9.59) with boundary conditions (9.61)–(9.64) was solved numerically.

In case of the equal sign of the topological charges of the C -points in the incident beams, the interference of the beams at $z=0$ gives one or three C -points (depending on d/w , h and θ). The sum of topological charges of all of the C -points in both cases is equal to $1/2$ or $-1/2$. The propagation of sole C -point is not of significant interest. The third C -point, which appears between two initially existing C -points, has the topological charge with the sign opposite to that of the initial C -points. Its coordinates $x_C = 0$, $y_C = -0.5d \operatorname{tn}(\theta/2)$ do not change with the propagation. After traversing a distance of $\tilde{z} = L_d [(d/w)^2 (1 + \cos\theta)^{-1} - 1]^{1/2}$ a dynamic inversion of its topological charge [81] takes place: two C -points with opposite topological charges annihilate, and afterwards another pair of the “oppositely charged” C -points appears immediately.

In the case of opposite topological charges of the C -points in the incident beams, there can be two or four C -points at $z=0$ as a result of their interference. The sum of topological charges of all these C -points is equal to zero. Let us consider a case, when there is one pair of C -points with opposite topological charges located in the centers of incident beams. If $\theta = 0$, then for any value of $z = z_1$ the light in each point of the straight line ($x = 0; z = z_1$) is right-handed circularly polarized, because the left-handed circularly polarized Laguerre–Gaussian components suppress each other at this line. In this case the above mentioned straight line is a symmetry axis for the intensity distribution in the transversal plane. If $\theta \neq 0$ and $d/w > \sqrt{2}$, then after traversing some distance additional pair of C -points with opposite topological charges is nucleated (Fig. 9.23a). Corresponding transversal polarization distribution for $z = 2L_d$ is shown in Fig. 9.23b. If the phase shift θ between the initial beams exceeds certain critical value $\theta_{cr}(d/w)$, then one of the “new” C -points will subsequently annihilate with one of the “old” C -points having the opposite topological charge (Fig. 9.23c). The transversal polarization distribution at $z = 2L_d$ for this case is shown in Fig. 9.23d. The function $\theta_{cr}(d/w)$ increases monotonically with d/w .

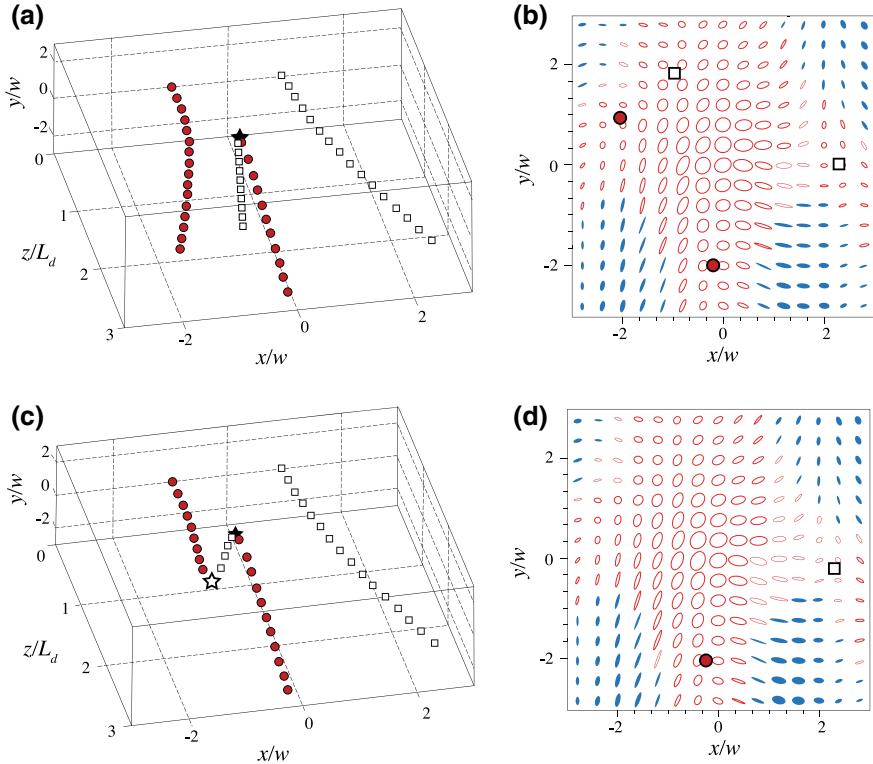


Fig. 9.23 (Adopted from [80]) The processes of the nucleation and the annihilation of the C-points (**a**, **c**) and the transversal polarization distributions at $z=3L_d$ (**b**, **d**). Here $\theta=20^\circ$ in (**a**, **b**), and $\theta=30^\circ$ in (**c**, **d**). The topological charges of the C-points in the initial (separate) beams are opposite, $d/w=2$, $h=1$. The critical phase shift for this case is $\theta_{cr} \approx 24^\circ$. Circles indicate the C-points with positive topological charge, and squares indicate those with the negative topological charge. Pairwise creation of the C-points is indicated by the filled stars in (**a**) and (**c**), and pairwise annihilation is indicated by the empty star in (**c**)

In a nonlinear medium the features of the interaction of the polarization singularities is determined by the parameters d/w , h , θ , σ_2/σ_1 , ρ_1/σ_1 and $P=\sigma_1 L_d E_0^2$. The detail numerical studies have shown its predominant influence on the propagation features of the C-points. With the growth of P these features take place for smaller values of the medium nonlinear susceptibility, or at smaller z . In case of the equal topological charges of the C-points the self-action promotes the amplification of the central and lateral intensity maxima and their stability during the propagation. If $\sigma_1/2 + \sigma_2 > 0$, the circularly polarized components A_\pm focus each other due to the nonlinear cross-interaction and each of the lateral maxima will be attracted by the corresponding Gaussian component with the opposite handedness of the polarization. As a result, there appear not three, but only two intense peaks in the transversal section of the propagating light, and their polarization is determined by the value of h .

In case of significant focusing nonlinearity the central spots of the Gaussian components do not merge and propagate separately instead.

If $h \approx 1$, the polarization is almost linear. The intensity of light in the vicinity of each of the C -points, where the radiation is right-handed circularly polarized tends to zero, since the power of this circularly polarized component is attracted by the component with the opposite rotation (right-handed circularly polarized Gaussian is attracted by the left-handed circularly polarized Laguerre–Gaussian component). If $\sigma_1/2 + \sigma_2 < 0$, then the intensity maxima with opposite handedness of the polarization rotation defocus each other and tend to occupy such areas in the beam cross-section that do not overlap. The polarization and the intensity distributions in this case become similar to those shown in Figs. 9.2 and 9.23 (for approximately the same incident beams parameters). The maxima of the intensity maintain their shapes during the propagation owing to the nonlinear self-action, and become more contrasting. The larger the absolute value of the negative quantity σ_2/σ_1 , the stronger is the difference between the polarization state of this maxima and linear polarization state.

For any value of $\sigma_1/2 + \sigma_2$ originating due to the interference of two beams having singularities with equal topological charges the central C -point does not move straight along the z -axis, as was the case in linear medium (Fig. 9.2a), and the dynamic inversion of its topological charge does not occur. Instead, there takes place two separate events of pairwise creation and annihilation of the C -points (Fig. 9.24). Depending on the sign of $\sigma_1/2 + \sigma_2$, the sequence of these two events

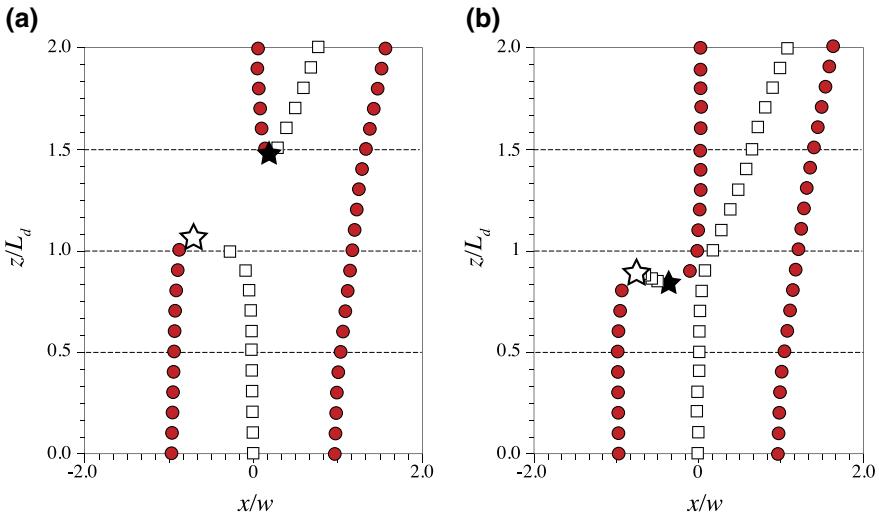


Fig. 9.24 (Adopted from [80]) C -lines in a nonlinear medium for $P = 1$, $\sigma_2/\sigma_1 = 0.5$, $\rho_1/\sigma_1 = 0$ (a) and $P = 2$, $\sigma_2/\sigma_1 = -0.8$, $\rho_1/\sigma_1 = 0$ (b). The parameters of the incident beams are $d/w = 2$, $\theta = 30^\circ$ and $h = 1$. Circles designate the C -lines with the topological charge $1/2$, while squares designate ones with the topological charge $-1/2$. Pairwise creation of the C -points is indicated by filled stars and pairwise annihilation is indicated by empty ones

can be different. Figure 9.25a, b shows the dependence of the nucleation/annihilation distance on the dimensionless power. As can be seen, the growth of P results in the increase of the z -coordinates of the C -point's nucleation/annihilation processes for $\sigma_1/2 + \sigma_2 > 0$ and in the decrease of these characteristic coordinates in the case of $\sigma_1/2 + \sigma_2 < 0$.

In case of incident beams with opposite topological charges of the polarization singularities, the trajectories of the C -points are similar to those shown in Fig. 9.23. Assume that the phase shift $\theta > \theta_{cr}$, so the pairwise annihilation of C -points is observed at certain $z = z^*$ in linear medium (see Fig. 9.23c). If $\sigma_1/2 + \sigma_2 > 0$ the evolution of the C -points proceeds slower, while in the case then $\sigma_1/2 + \sigma_2 < 0$ it proceeds faster, comparing with the propagation in a linear medium. Such behavior is typical for small values of the dimensionless power P .

For large values of P the self-action of the circularly polarized components A_{\pm} prevents the interaction of the incident beams and, therefore, the interaction of the C -points in them. The z -coordinate of the pairwise creation of C -points is affected by the change of the dimensionless power in a similar way (Fig. 9.25c, d).

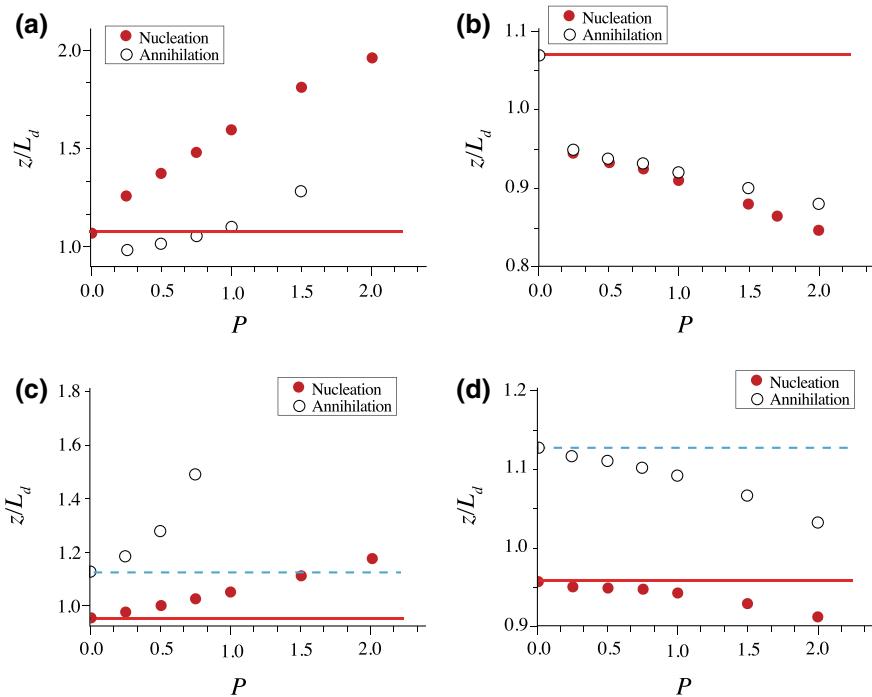


Fig. 9.25 (Adopted from [80]) The dependence of the distance from the border of the medium, where the nucleation/annihilation of two C -points takes place, on the dimensionless power parameter. Filled circles correspond to the nucleation and empty circles correspond to annihilation. Equal topological charges of the initial singularities (a, b), $\theta = 30^\circ$. Opposite topological charges of the initial singularities (c, d), $\theta = 40^\circ$. Other parameters: $d/w = 2$, $h = 1$, $\rho_1 = 0$, $\sigma_2/\sigma_1 = 0.5$ (a, c), -0.8 (b, d)

The weak spatial dispersion of the cubic nonlinearity enhances the self-action of one of the circularly polarized components of propagating light and insignificantly causes on pairwise creation process. The growth of spatial dispersion influence leads to the formation of strongly inhomogeneous polarization distribution in beam cross-section.

9.7 Nonlinear Optical Activity in Propagation of Ultrashort Elliptically Polarized Laser Pulses in a Medium with Frequency Dispersion and Significant Spatial Dispersion of Cubic Nonlinearity

9.7.1 *Physical Model of Third-Order Nonlinear Optical Susceptibility of Isotropic Medium with Frequency Dispersion and Significant Spatial Dispersion of Cubic Nonlinearity*

In the present section, we describe the model of the nonlocal nonlinear optical response of the isotropic gyrotropic medium (medium symmetry group $\infty\infty$) with frequency dispersion, allowing one to formulate the material equations without the requirement of smallness of characteristic dimension scale of the nonlocality and without the limitations on the duration of the propagating pulse [82]. This model gives an opportunity to use the modification of the finite-difference time-domain method with the auxiliary differential equation for the description of the propagation of elliptically polarized pulses of arbitrary duration in such a medium. For long pulses the results of the numerical analysis coincide with those obtained analogously within the slowly varying envelope approximation. For pulses containing about 10 or less oscillations of the electric field vector the results of numerical simulations significantly differ from those predicted by the slowly varying envelope approximation, particularly, from the analytic expressions for the intensity-dependent ellipticity degree and the angle of rotation of the polarization ellipse obtained in the case of propagation of monochromatic light.

When considering the plane electromagnetic wave (one-dimensional case) propagating along the z axis, expressions (9.9) becomes simpler

$$\mathbf{P}^{(1)}(z, t) = \int_0^\infty dt_1 \int_{-\infty}^\infty dz_1 \hat{\chi}^{(1)}(t_1, z, z_1) \mathbf{E}(z_1, t - t_1), \quad (9.65)$$

$$\begin{aligned} \mathbf{P}^{(3)}(z, t) = & \int_0^\infty dt_1 \int_0^\infty dt_2 \int_0^\infty dt_3 \int_{-\infty}^\infty dz_1 \int_{-\infty}^\infty dz_2 \int_{-\infty}^\infty dz_3 \hat{\chi}^{(3)}(t_1, t_2, t_3, z, z_1, z_2, z_3) : \mathbf{E}(z_1, t - t_1) \\ & \mathbf{E}(z_2, t - t_2) \mathbf{E}(z_3, t - t_3). \end{aligned} \quad (9.66)$$

If we assume the medium to be in a half-space $z > 0$, when the following form of the tensor will satisfy to all mentioned in Sect. 9.3.2 requirements for $z_1 > 0$:

$$\chi_{ij}^{(1)}(t_1, z, z_1) = g^{(1)}(t_1) [\delta_{ij} + \gamma_1(\delta_{xi}\delta_{yj} - \delta_{xj}\delta_{yi})(z - z_1)] \exp[-(z - z_1)^2/d_1^2]/(\pi d_1). \quad (9.67)$$

In (9.67) d_1 and γ_1 are the parameters characterizing linear gyrotropy of the medium, here and further $i, j = x, y$. When the frequency dispersion is the Lorentz type [83], then $g^{(1)}(t) = (1/4\pi)\{(\epsilon_\infty - 1)\delta(t_1) + \omega_0^2 - (\epsilon_s - \epsilon_\infty)(\omega_0^2 - \delta_0^2)^{-1/2} \exp(-\delta_0 t_1) \sin[(\omega_0^2 - \delta_0^2)^{1/2} t_1]\}$, ϵ_s , δ_0 , ϵ_∞ , ω_0 are constants. For $z_1 < 0$, $\chi_{ij}^{(1)}(t_1, z, z_1) \equiv 0$. The substantiation of choice of such a form of $\chi_{ij}^{(1)}(t_1, z, z_1)$ is given in [28], where the expressions (9.65) and (9.67) were used for the study of the propagation of ultrashort (about several oscillations of the light field) laser pulses in a linear medium with both frequency and spatial dispersion by means of integration of Maxwell equations by the finite-difference time-domain method with the auxiliary differential equation. The reasonable generalization of this model for the nonlinear isotropic gyrotropic medium with frequency dispersion and spatial dispersion of cubic nonlinearity could be the following ($z_{1,2,3} > 0$):

$$\begin{aligned} \chi_{ijkl}^{(3)}(t_1, t_2, t_3, z, z_1, z_2, z_3) = & a\delta(t_1)\delta(t_2)\delta(t_3)\delta(z_1)\delta(z_2)\delta(z_3)[\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}] \\ & + \delta(t_1)\delta(t_2 - t_3)g_3(t_3)\delta(z - z_1)\delta(z_2 - z_3)\left(\exp[-(z - z_3)^2/d_3^2]/(\sqrt{\pi}d_3)\right) \\ & \times [b\delta_{ij}\delta_{kl} + c(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) + \gamma_3(z - z_3)(\delta_{xi}\delta_{yj} - \delta_{yi}\delta_{xj})\delta_{kl}] \\ & + \delta(t_2)\delta(t_1 - t_3)g_3(t_1)\delta(z - z_2)\delta(z_1 - z_3)\left(\exp[-(z - z_1)^2/d_3^2]/(\sqrt{\pi}d_3)\right) \\ & \times [b\delta_{ik}\delta_{jl} + c(\delta_{ij}\delta_{kl} + \delta_{il}\delta_{jk}) + \gamma_3(z - z_1)(\delta_{xi}\delta_{yk} - \delta_{yi}\delta_{xk})\delta_{jl}] \\ & + \delta(t_3)\delta(t_1 - t_2)g_3(t_2)\delta(z - z_3)\delta(z_1 - z_2)\left(\exp[-(z - z_2)^2/d_3^2]/(\sqrt{\pi}d_3)\right) \\ & \times [b\delta_{il}\delta_{jk} + c(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl}) + \gamma_3(z - z_2)(\delta_{xi}\delta_{yl} - \delta_{yi}\delta_{xl})\delta_{jk}]. \end{aligned} \quad (9.68)$$

For $z_{1,2,3} < 0$, $\chi_{ijkl}^{(3)}(t_1, t_2, t_3, z, z_1, z_2, z_3) \equiv 0$. Here a , b , and c determine the cubic nonlinearity of the medium, γ_3 and d_3 determine its spatial dispersion. Indices i, j, k and l take values of x and y . Exponential choice of the coordinate dependence is prescribed only by the requirement of rapid decay of $\chi_{ijkl}^{(3)}(t_1, t_2, t_3, z, z_1, z_2, z_3)$ to zero with the increase of $|z - z_{1,2,3}|$. Terms $g_3(t_{1,2,3})$ we take as $g_3(\tilde{t}) = (\tau_1^2 + \tau_2^2)(\tau_1 \tau_2^2)^{-1} \exp(-\tilde{t}/\tau_2) \sin(\tilde{t}/\tau_1)$, which was thoroughly substantiated in [33]

and references therein. In last expression $\tau_{1,2}$ are the relaxation time constants, b/a and c/a determine the relative contributions of Kerr-type and Raman-type.

This model does not take into account the fact that the response of the homogeneous medium with sharp boundary will be different near the surface of the medium and in its bulk. It does not account for a number of effects, which are described by a Debye model. However, we prefer to use the model, which represents itself the generalization (for the case of spatial dispersion) of a widely used and efficiently working dispersion model. Below we do not take into account second-order nonlinear processes, which may result in considerable modification of the frequency and spatial spectrum of the propagating tightly focused ultrashort light pulse with broad frequency and spatial spectra. It would be correct to use (9.65)–(9.68) under conditions when the effective quadratic optical susceptibility is negligibly small, or when the phase-matching condition cannot be satisfied.

9.7.2 Propagation of Ultrashort Elliptically Polarized Pulses in a Medium with Frequency Dispersion and Nonlocality of Nonlinear Optical Response

Let us consider elliptically polarized light pulse with the Gaussian envelope, which propagates in vacuum towards the plane boundary $z=0$ of the medium. We assume that at $t=0$ the Cartesian components of the electric field in the pulse do not depend on x and y , and they are expressed by the following formulas:

$$E_x(z, t=0) = \left[(\tilde{P}I_0/2) \left(1 - \sqrt{1 - M_0^2} \right) \right]^{1/2} \exp\left(-\frac{(z-z_0)^2}{w_0^2} \right) \text{sign}(M_0) \sin\left(\frac{2\pi(z-z_0)}{\lambda} \right), \quad (9.69)$$

$$E_y(z, t=0) = \left[(\tilde{P}I_0/2) \left(1 + \sqrt{1 - M_0^2} \right) \right]^{1/2} \exp\left(-\frac{(z-z_0)^2}{w_0^2} \right) \sin\left(\frac{2\pi(z-z_0)}{\lambda} \right). \quad (9.70)$$

Here λ is wavelength. The dimensionless intensity $I = (E_x^2 + E_y^2)/I_0$ achieves its maximum value \tilde{P} in $z=z_0$. The incident pulse has the ellipticity degree of the polarization ellipse, which is equal to M_0 , and the main axis of the polarization ellipse is parallel to the y axis (we always can choose the coordinate axes in such a way for a medium with $\infty\infty$ symmetry).

For the pulses containing about ten or less oscillations of the electric field the results of numerical simulations significantly differ [82] from the analytic expressions for the intensity-dependent ellipticity degree and the angle of rotation of the polarization ellipse in the case of propagation of monochromatic radiation predicted

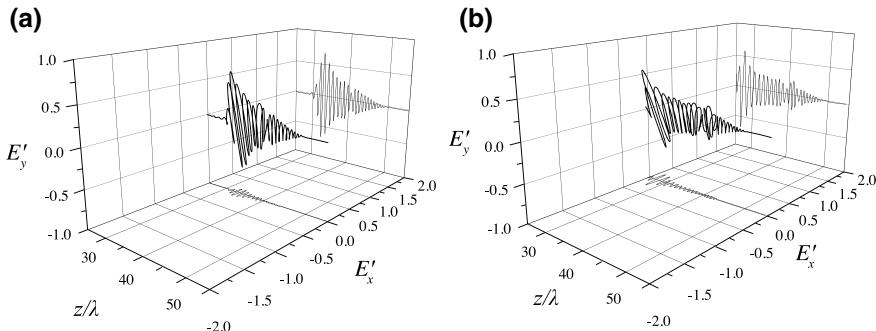


Fig. 9.26 (Adopted from [82]) The hodograph of the electric field strength vector for the pulse traversed 40 wavelengths in the nonlinear medium, $\tilde{P} = 2.5$ (a) and $\tilde{P} = 5$ (b). The half width of the incident pulse is five wavelengths. The nonlinear medium parameters and the parameters of the incident pulse are $\gamma_3 = -0.1(\lambda I_0)^{-1}$, $d_3 = 0.05\lambda$, $a = 0$, $b = c = 10^{-3}I_0^{-1}$, the other parameters are $\varepsilon_s = 5.25$, $\varepsilon_\infty = 2.25$, $\gamma_1 = 0$, $d_1 = 0$, $\delta_0 = 1.64 \cdot 10^{-5}\omega$, $\omega_0 = 0.46\omega$, $\tau_1 = 10.5/\omega$, $\tau_2 = 27.6/\omega$, $w_0 = 100\lambda$, $M_0 = 0.1$

by the slowly varying envelope approximation. Owing to the nonlinearity of the medium and to the nonlocality of its linear and nonlinear response, the propagating pulse, which was linearly polarized at the medium border acquires an orthogonal polarization component during the propagation. A hodograph (a curve in the space of parameters $E'_x = E_x/(PI_0)^{1/2}$, $E'_y = E_y/(PI_0)^{1/2}$ and z/λ , traced by the end of the electric field vector) is shown in Fig. 9.26 for the pulse with initial half width of five wavelengths after traversing in a medium 40 wavelengths with $\tilde{P} = 2.5$ (a) and $\tilde{P} = 5$ (b). The line projected to the plane $E'_x = \text{const}$ represents the dependence $E'_y(z/\lambda)$, and the line projected to the plane $E'_y = \text{const}$ represents the dependence $E'_x(z/\lambda)$. The increase of the initial intensity results in the increase of the orthogonal polarization component. The hodograph looks like a deformed helix with changing radius, and its axis coincides with the z axis. In this case it is not possible to discuss the polarization state of light.

In a number of cases the helicity of the hodograph changes along the pulse (for example, from right to left, or vice versa), which reflects the change of the sense of rotation of the electric field vector. Such a case is well illustrated in Fig. 9.3b, where the fragmentation of the pulse into two parts occurred, and the senses of rotation of the electric field vector within these parts are opposite.

In the medium without spatial dispersion the hodographs of the electric field vector of pulses with initial values of ellipticity degree $-M_0$ and M_0 are absolutely reflection symmetric with respect to one another relatively to the plane yz (the optical properties of such a medium are identical for right- and left-handed circularly polarized components of light). In the presence of nonlocality of nonlinear

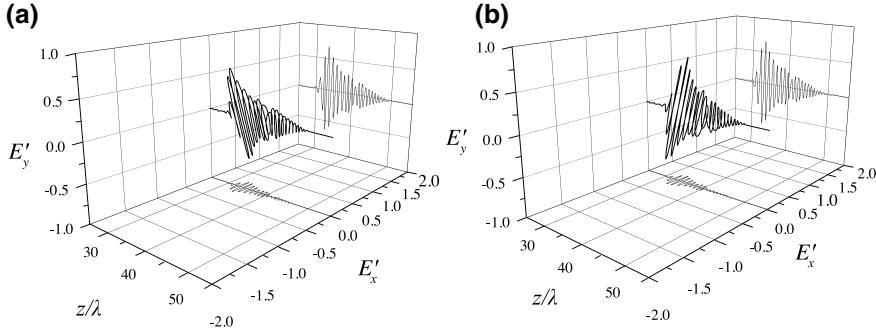


Fig. 9.27 (Adopted from [82]) The hodograph of the electric field strength vector for the pulse with peak intensity $\tilde{P}=1$, half width of five wavelengths and the ellipticity degree $M_0=0.25$ **a** and $M_0=-0.25$ **b** after traversing 40 wavelengths in a medium with spatial dispersion of nonlinearity; $b=c=2.5 \cdot 10^{-3} I_0^{-1}$, $\gamma_3=0$, $d_3=0$, the other parameters are the same as in Fig. 9.26

optical response the hodographs of two pulses with initial ellipticities M_0 and $-M_0$ may rotate either in the same direction or in opposite directions with different rates of rotation (see Fig. 9.27).

Under certain conditions there can be realized the situation, when the hodograph of one pulse is rotating, and the hodograph of another almost does not rotate. The numerical investigations show that the pulse possessing initially circular polarization maintains its polarization state during the propagation. The same result is given by the slowly varying envelopes method in a medium with spatial dispersion of cubic nonlinearity. The nonzero value of γ_3 provides the intensity-dependent difference of the absorption lengths for the circularly polarized components of the light field with opposite rotation directions (nonlinear circular dichroism), also predicted by the slowly varying envelopes method. The latter is illustrated by Fig. 9.28,

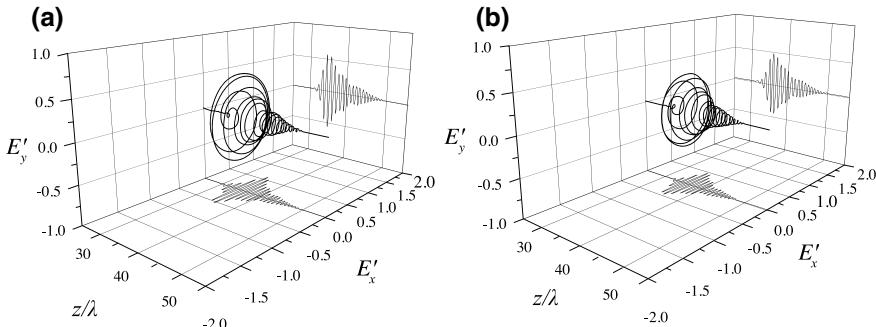


Fig. 9.28 (Adopted from [82]) The hodograph of the electric field strength vector for the circularly polarized pulse with peak intensity $\tilde{P}=2.5$, half width of five wavelengths, and the ellipticity degree $M_0=1$ **(a)** and $M_0=-1$ **(b)**, after traversing 40 wavelengths in a medium with spatial dispersion of nonlinearity. The medium parameters are the same as in Fig. 9.26

where the hodographs of the electric field vectors for pulses with $M_0 = 1$ (a) and $M_0 = -1$ (b) are shown for a distance of 40 wavelengths. The absorption for the pulse with $M_0 = -1$ (b) is remarkably stronger than that for a pulse with $M_0 = 1$.

9.8 Conclusions

In the middle of 70's the first experimental evidence of the nonlinear optical activity came to light, giving the impulse to the development of nonlinear polarization optics. The subsequent theoretical and experimental investigations assure that the polarization self-action and interaction of waves are fine and widespread phenomena in nonlinear optics. Owing to the appearance of metamaterials in optics, many fascinating phenomena which earlier required complicated experimental setups to be detected, now can be readily reproduced and studied. Many different experiments stimulate theoretical investigations of laser light propagation in metamaterials. In particular, recently giant nonlinear optical activity was experimentally observed in a plasmonic metamaterial [84]. It exhibits nonlinear optical activity 30 million times stronger than lithium iodate crystal, the first one where nonlinear optical activity was observed about forty years ago [41, 42].

Now nonlinear polarization optics begins rapid development, due to the requests of high technology. For example, the interest in metamaterials composed of a two dimensional periodic array of three-dimensional spirals is primarily related to the possibility of using them in compact elements for controlling the polarization of ultrashort laser pulses (see [85] and references therein). These structures, characterized by significantly different transmittances for circularly polarized pulses with opposite directions of rotation of electric-field strength vectors, are promising [85] for generators of circularly polarized light, which can be used in various applications [81, 86]. The methods for generating circularly polarized light based on the use of quarter-wave plates or cholesteric liquid crystals with a very small mismatch between the spiral pitch and the wavelength of propagating wave, which are conventional for monochromatic waves, cannot be applied for ultrashort laser pulses because of the wide spectrum of the latter. Only such methods as the finite-difference time-domain method are promising for analysis of the effect of the parameters of structural cell of nonlinear metamaterial on the transmission and reflection of elliptically polarized long and ultrashort pulses incident on a metamaterial sample [87, 88]. Effective spatial dispersion of these media cannot be considered as small and it is very difficult (in some cases impossible) to obtain analytical expressions for the intensity, the ellipticity degree of polarization ellipse, and the inclination angle of its major axis in this case. This is the reason, why these lines may be considered as the end of one of final chapters of nonlinear optics, containing beautiful analytical formulae describing the propagation of electromagnetic fields in space.

Acknowledgements The author takes this opportunity to acknowledge many key contributions to this field of nonlinear polarization optics made by his former students and postgraduate students at M.V. Lomonosov Moscow State University. I am indeed grateful to Prof. Dr. A.A. Golubkov, Dr. S.N. Volkov, Dr. I.A. Perezhogin, Dr. N.N. Potravkin, Dr. K.S. Grigoriev, and M.P.S. G.A. Gryaznov. Some results of common work are observe in this chapter. I have endeavored to provide within the practical limitations of a chapter as comprehensive a review of theoretical problems of singular nonlinear polarization optics.

References

1. S.N. Bagaev, O.N. Krokhin, A.A. Manenkov, Pages from the history of quantum electronic research in the Soviet Union. *J. Mod. Opt.* **52**, 1657–1669 (2005)
2. V.A. Makarov, Quantum electronics and the R.V. Khokhlov–S.A. Akhmanov school of coherent and nonlinear optics at Moscow State University. *Phys. Usp.* **47**, 1059–1065 (2004)
3. V.G. Dmitriev, L.V. Tarasov, *Prikladnaya Nelineinaya Optika (Applied Nonlinear Optics)* (Fizmatlit, Moscow, 2004)
4. S.V. Chekalin, V.P. Kandidov, From self-focusing light beams to femtosecond laser pulse filamentation. *Phys. Usp.* **56**, 123–140 (2013)
5. V.A. Makarov, Nonlinear optics: its background, present, and future. *Herald Russ. Acad. Sci.* **81**(3), 238–245 (2011)
6. S.I. Vavilov, Microstructure of light; investigations and studies. *USSR Academy of Science* (1950)
7. P.D. Maker, R.W. Terhune, S.M. Savage, Intensity-dependent changes in the refractive index of liquids. *Phys. Rev. Lett.* **12**, 507–509 (1964)
8. S.A. Akhmanov, V.I. Zharikov, Nonlinear optics of gyrotropic media. *JETP Lett.* **6**, 137–140 (1967)
9. S.N. Volkov, V.A. Makarov, I.A. Perezhogin, Formation of inhomogeneously polarized light beams at the sum frequency by two collinear elliptically polarized Gaussian beams focused into chiral medium. *Quantum Electron.* **36**, 860–866 (2006)
10. V.A. Makarov, I.A. Perezhogin, Generation of reflected second harmonic light beam with inhomogeneous transversal distribution of polarization from the surface of chiral medium by normally incident Gaussian beam. *Opt. Commun.* **281**, 3906–3912 (2008)
11. A.A. Golubkov, V.A. Makarov, I.A. Perezhogin, Formation of elliptically polarized ring-shaped structures of electric field during light self-focusing in isotropic medium with spatial dispersion of nonlinearity. *Moscow University Physics Bulletin*, No 1, pp. 54–57 (2009)
12. V.A. Makarov, I.A. Perezhogin, Transversal structure of a sum-frequency beam generated from the surface of a chiral medium, *J. Opt. A: Pure Appl. Opt.* **11**, 074008(7) (2009)
13. V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Specific features of the self-action of elliptically polarized light pulses and the formation of vector solitons in an isotropic medium with the anomalous frequency dispersion and the spatial dispersion of cubic nonlinearity. *Laser Phys.* **19**, 322–329 (2009)
14. R. Azaam, N. Bashara, *Ellipsometry and Polarised Light* (Amsterdam, North Holland, 1977)
15. Y.R. Shen, *The Principles of Nonlinear Optics* (Wiley-Interscience, 2003)
16. J.F. Nye, Lines of circular polarization in electromagnetic wave fields. *Proc. R. Soc. Lond. A* **389**, 279–290 (1983)
17. M.R. Dennis, Polarization singularities in paraxial vector fields: morphology and statistics. *Opt. Commun.* **213**, 201–221 (2002)
18. M.V. Berry, M.R. Dennis, Polarization singularities in isotropic random waves. *Proc. R. Soc. A* **457**, 141–155 (2001)

19. M.V. Berry, M.R. Dennis, The optical singularities of birefringent dichroic chiral crystals. Proc. R. Soc. A **459**, 1261–1292 (2003)
20. Y. Egorov, T. Fadeyeva, A. Volyar, The fine structure of singular beams in crystals: colours and polarization. J. Opt. A Pure Appl. Opt. **6**, 217–228 (2004)
21. F. Flossmann, T. Ulrich, U.T. Schwarz, M. Maier, M.R. Dennis, Polarization singularities from unfolding an optical vortex through a birefringent crystal. Phys. Rev. Lett. **95**, 253901 (4) (2005)
22. G.V. Bogatyryova, K.P. Felde, P.V. Polyanskii, M.S. Soskin, Nongeneric polarization singularities in combine vortex beams. Opt. Spectrosc. **97**, 782–789 (2004)
23. M.R. Dennis, Polarization singularity anisotropy: determining monostardom. Opt. Lett. **33**, 2572–2574 (2008)
24. O.V. Angelsky, I.I. Mokhun, A.I. Mokhun, M.S. Soskin, Interferometric methods in diagnostics of polarization singularities. Phys. Rev. E **65** 036602 (5) (2002)
25. Y.F. Chen, T.H. Lu, K.F. Huang, Visualization of vector singularities. Phys. Rev. Lett. **96**, 033901 (4) (2006)
26. K. O'Holleran, F. Flossmann, M.R. Dennis, M.J. Padgett, Methodology for imaging the 3D structure of singularities in scalar and vector optical fields. J. Opt. A: Pure Appl. Opt. **11**, 094020–094026 (2009)
27. M.R. Dennis, Polarization singularities in paraxial vector fields: morphology and statistics. Opt. Commun. **213**, 201–221 (2002)
28. N.N. Potravkin, I.A. Perezhogin, V.A. Makarov, Numerical solution of Maxwell equations by a finite-difference time-domain method in a medium with frequency and spatial dispersion. Phys. Rev. E **86**, 056706 (6) (2012)
29. M.P. Silverman, Reflection and refraction at the surface of a chiral medium: comparison of gyrotropic constitutive relations, invariant or noninvariant under a duality transformation. J. Opt. Soc. Am. A **6**, 830–837 (1986)
30. L.D. Landau, E.M. Lifshitz, E.M. Pitaevskii, *Electrodynamics of Continuous Media*, 2nd edn. (Pergamon Press, Oxford, 1984)
31. V.M. Agranovich, V.L. Ginzburg, *Crystal Optics with Spatial Dispersion and Excitons* (Springer, 1984)
32. Y.I. Sirotin, M.P. Shaskol'skaya, *Fundamentals of Crystal Physics* (Mir Publishers, Moscow, 1982)
33. S.V. Popov, Y.P. Svirko, N.I. Zheludev, Susceptibility tensors for nonlinear optics. *Optics and Optoelectronics Series*, Ed. by E.R. Pike, B.E.A. Saleh, S. Lowenthal, W.T. Welford (Institute of Physics Publishing, Bristol and Philadelphia, 1995)
34. J.F. Nye, *Physical Properties of Crystals* (Clarendon Press, Oxford, 1964)
35. A.A. Golubkov, V.A. Makarov, Boundary conditions for electromagnetic field on the surface of media with weak spatial dispersion. Phys. Usp. **38**, 325–332 (1995)
36. A.A. Golubkov, V.A. Makarov, Boundary conditions for electromagnetic field on the surface of linear and nonlinear crystals: allowance for weak spatial dispersion and near surface inhomogeneity of optical properties at the intermedium boundary. J. Russ. Laser Res. **17**, 480–487 (1996)
37. A.A. Golubkov, V.A. Makarov, Material equation for the polarization current on the surface of media with weak spatial dispersion. Laser Phys. **6**, 1015–1020 (1996)
38. A.A. Golubkov, V.A. Makarov, Spectroscopy of nonlinear gyrotropic medium and surface diagnostics based on polarization effects due to self-action of light. J. Mod. Opt. **37**, 1531–1543 (1990)
39. S.A. Akhmanov, G.A. Lyakhov, V.A. Makarov, V.I. Zharikov, Theory of nonlinear optical activity in isotropic media and liquid-crystals. Optica Acta **29**, 1359–1369 (1982)
40. N.I. Zheludev, A.D. Petrenko, G.I. Trush, Nonlinear reflection optical activity: nonlinear refraction anisotropy. Sov. Phys. Crystallogr. **32**, 399–405 (1987)
41. A.I. Kovrigin, D.V. Yakovlev, B.V. Zhdanov, N.I. Zheludev, Self-induced optical activity in crystals. Opt. Commun. **35**, 92–95 (1980)

42. S.A. Akhmanov, B.V. Zdanov, N.I. Zheludev, A.I. Kovrigin, V.I. Kuznetsov, Nonlinear optical activity in crystals. JETP Lett. **29**, 294–298 (1979)
43. A.D. Petrenko, N.I. Zheludev, Physical mechanisms of nonlinear optical activity in crystals. Optica Acta **31**, 1177–1184 (1984)
44. A.A. Golubkov, V.A. Makarov, Polarization spectroscopy of nonlinear rotation and deformation of the light polarization ellipse transmitted through nonlinear gyrotropic crystals. Mosc. Univ. Phys. Bull. **44**(2), 58–63 (1989)
45. A.A. Golubkov, V.A. Makarov, Spectroscopy of nonlinear optical rotation and deformation of the polarization ellipse for light reflected from nonlinear gyrotropic crystals. Opt. Spectrosc. **67**, 669–672 (1989)
46. A.A. Golubkov, V.A. Makarov, Two-wave spectroscopy of nonlinear optical rotation and deformation of the polarization ellipse of light reflected from a nonlinear gyrotropic crystal. Opt. Spectrosc. **69**, 369–372 (1990)
47. S.N. Volkov, N.I. Koroteev, V.A. Makarov, Second-harmonic generation in the interior of an isotropic medium with quadratic nonlinearity by a focused inhomogeneously polarized pump beam. JETP **86**, 687–695 (1998)
48. A.P. Shkurinov, A.V. Dubrovskii, N.I. Koroteev, Second harmonic generation in an optically active liquid: Experimental observation of a fourth-order optical nonlinearity due to molecular chirality. Phys. Rev. Lett. **70**, 1085–1088 (1993)
49. A.V. Dubrovskii, N.I. Koroteev, A.P. Shkurinov, Experimental observation of a five wave nonlinear optical process in an optically active liquid: a second harmonic generation which is sensitive to the mirror asymmetry of biomolecules. JETP Lett. **56**, 570–574 (1992)
50. N.I. Koroteev, BioCARS—a novel nonlinear optical technique to study vibrational spectra of chiral biological molecules in solution. Biospectroscopy **1**, 341–350 (1995)
51. N.I. Koroteev, Novel nonlinear optical techniques for studying chiral molecules of biological importance, in *Frontiers in Nonlinear Optics. The Sergei Akhmanov Memorial Volume*, ed. by H. Walther, N.I. Koroteev, M. Scully (Institute of Physics Publishing, Bristol, England, 1993), pp. 228–239
52. N.I. Koroteev, New schemes for nonlinear optical spectroscopy of solutions of chiral biological macromolecules. JETP **79**, 681–690 (1994)
53. V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Polarisation singularities in the electric field at a sum-frequency generated by two collinear elliptically polarised Gaussian beams in the bulk of a nonlinear gyrotropic medium. Quantum Electron. **41**, 149–152 (2011)
54. N.I. Koroteev, V.A. Makarov, S.N. Volkov, Sum-frequency generation in the bulk of an isotropic gyrotropic medium with two collinear pump beams. Laser Phys. **9**, 655–664 (1999)
55. K.S. Grigoriev, V.A. Makarov, I.A. Perezhogin, Polarization singularities in a sum-frequency light beam generated by a bichromatic singular beam in the bulk of an isotropic nonlinear chiral medium. Phys. Rev. A **92**, 023814 (7) (2015)
56. V.A. Makarov, I.A. Perezhogin, Nonuniform transverse distribution of the light intensity and polarization upon sum-frequency generation from the surface of an isotropic gyrotropic medium in the case of normal incidence of light. Quantum Electron. **39**, 627–633 (2009)
57. J.J. Maki, M. Kauranen, A. Persoons, Surface second-harmonic generation from chiral materials. Phys. Rev. B **51**, 1425–1434 (1995)
58. M. Kauranen, T. Verbiest, A. Persoons, Electric and magnetic contributions to the second-order optical activity of chiral surfaces. Nonlinear Opt. Princ. Mater. Phenom. Devices **8**, 243–249 (1994)
59. S.V. Elshocht, T. Verbiest, M. Kauranen, A. Persoons, B.M.W. Langeveld-Voss, E.W. Meijer, Direct evidence of the failure of electric-dipole approximation in second-harmonic generation from a chiral polymer. J. Chem. Phys. **107**, 8201–8203 (1997)
60. M. Kauranen, J.J. Maki, T. Verbiest, S. Van Elshocht, A. Persoons, Quantitative determination of electric and magnetic second-order susceptibility tensors of chiral surfaces. Phys. Rev. B **55**, R1985–R1988 (1997)

61. S.N. Volkov, N.I. Koroteev, V.A. Makarov, Second harmonic generation by reflection of a two-dimensional laser beam from the surface of a isotropic gyrotropic medium. *Quantum Electron.* **27**, 517–522 (1997)
62. S.N. Volkov, N.I. Koroteev, V.A. Makarov, Second harmonic generation from a surface of an isotropic and gyrotropic medium in the case of normal incidence of a Gaussian pumping beam. *Opt. Spektrosk.* **85**, 286–292 (1998)
63. R. Stolle, M. Loddoch, G. Marowsky, Theory of second-harmonic circular dichroism at surfaces. *Nonlinear Opt. Princ. Mater. Phenom. Devices* **8**, 79–85 (1994)
64. Y.R. Shen, Surface contribution versus bulk contribution in surface nonlinear optical spectroscopy. *Appl. Phys. B* **68**, 295–300 (1999)
65. M.A. Kriech, J.C. Conboy, Label free chiral detection of melittin binding to planar supported lipid bilayer by chiral second harmonic generation. *Anal. Chim. Acta* **496**, 143–153 (2003)
66. M.J. Huttunen, M. Erkintalo, M. Kauranen: Absolute nonlinear optical probes of surface chirality. *J. Opt. A: Pure Appl. Opt.* **11**, 034006 (6) (2009)
67. K.S. Grigoriev, V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Singularities in the second harmonic light field polarisation arising upon reflection of normally incident elliptically polarised Gaussian beam from the surface of an isotropic chiral medium. *Quantum Electron.* **41**, 993–996 (2011)
68. V.A. Makarov, I.A. Perezhogin: Transversal structure of a sum-frequency beam generated from the surface of a chiral medium. *J. Opt. A: Pure Appl. Opt.* **11**, 074008 (11) (2009)
69. N.I. Koroteev, V.A. Makarov, S.N. Volkov, Sum frequency generation by reflection of light from the surface of chiral medium. *Nonlinear Opt. Princ. Mater. Phenom. Devices* **17**, 247–269 (1997)
70. V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Polarization singularities in second harmonic beam generated from the surface of the medium with spatial dispersion of nonlinear response. *J. Opt.* **14**, 055202 (10 pp) (2012)
71. K.S. Grigoriev, V.A. Makarov, I.A. Perezhogin, Formation of the lines of circular polarization in a second harmonic beam generated from the surface of an isotropic medium with nonlocal nonlinear response in the case of normal incidence. *J. Opt.* **18**, 014004 (8 pp) (2016)
72. A. Dreischuh, D.N. Neshev, V.Z. Kolev, S. Saltiel, M. Samoc, W. Krolikowski, Yu. Kivshar, Nonlinear dynamics of two-color optical vortices in lithium niobate crystals. *Opt. Express* **16**, 5406–5420 (2008)
73. S.A. Akhmanov, A.P. Sukhorukov, R.V. Khokhlov, Self-focusing and self-trapping of intense light beams in a nonlinear medium. *Phys. JETP* **23**, 1025–1037 (1966)
74. S.A. Akhmanov, A.P. Sukhorukov, R.V. Khokhlov, Self-focusing and diffraction of light in a nonlinear medium. *Phys. Usp.* **93**, 609–636 (1964)
75. S.A. Akhmanov, M.A. Vorontsov, V.P. Kandidov, A.P. Sukhorukov, S.S. Chesnokov, Thermal self-action of light beams and methods of compensating for it. *Radiophys. Quantum Electron.* **23**, 1–30 (1980)
76. A.A. Golubkov, V.A. Makarov, Amplitude and polarization effects in self-focusing of laser radiation in media with spatial dispersion of nonlinearity. *Radiophys. Quantum Electron.* **31**, 737–745 (1988)
77. S.N. Vlasov, V.A. Petrishchev, V.I. Talanov, Averaged description of wave beams in linear and nonlinear media (the method of moments). *Radiophys. Quantum Electron.* **14**, 1062–1070 (1971)
78. N.A. Panov, V.A. Makarov, K.S. Grigoriev, M.S. Yatskevitch, O.G. Kosareva, Generation of polarization singularities in the self-focusing of an elliptically polarized laser beam in an isotropic Kerr medium. *Physica D* **332**, 73–78 (2016)
79. K.S. Grigoriev, V.A. Makarov, G.M. Shishkov, Polarization singularities in the self-focusing of an elliptically polarized laser beam in an isotropic phase of nematic liquid crystal close to the temperature of phase transition. *Mol. Cryst. Liq. Cryst.* **650**, 23–31 (2017)
80. K.S. Grigoriev, V.A. Makarov, I.A. Perezhogin, Interaction of laser-generated polarization singularities in nonlinear isotropic gyrotropic medium. *J. Opt.* **16**, 105201 (8 pp) (2014)

81. T. Yoshioka, T. Ogata, T. Nonaka, M. Moritsugu, S.N. Kim, S. Kurihara, Reversible photon mode full color display by means of photochemical modulation of a helically cholesteric structure. *Adv. Mater.* **17**, 1226–1229 (2005)
82. G.A. Gryaznov, V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Modeling of nonlinear optical activity in propagation of ultrashort elliptically polarized laser pulses. *Phys. Rev. E* **89**, 013306 (6) (2014)
83. R.M. Joseph, S.C. Hagness, A. Taflove, Direct time integration of Maxwell's equations in linear dispersive media with absorption for scattering and propagation of femtosecond electromagnetic pulses. *Opt. Lett.* **16**, 1412–1414 (1991)
84. M. Ren, E. Plum, J. Xu, N.I. Zheludev, Giant nonlinear optical activity in a plasmonic metamaterial. *Nat. Commun.* **3**, 833–837 (2012)
85. J.K. Gansel, M. Thiel, M.S. Rill, M. Decker, K. Bade, V. Saile, G. von Freymann, S. Linden, M. Wegener, Gold helix photonic metamaterial as broadband circular polarizer. *Science* **325** (5947), 1513–1515 (2009)
86. G. de Filpo, F.P. Nicoletta, G. Chidichimo, Cholesteric emulsions for colored displays. *Adv. Mater.* **17**, 1150–1152 (2005)
87. V.A. Makarov, I.A. Perezhogin, N.N. Potravkin, Interaction of ultrashort elliptically polarized laser pulses with nonlinear helical photonic metamaterial. *Phys. Wave Phenom.* **23**, 14–20 (2015)
88. N.N. Potravkin, E.B. Cherepetskaya, I.A. Perezhogin, V.A. Makarov, Ultrashort elliptically polarized laser pulse interaction with helical photonic metamaterial. *Opt. Mater. Express* **4**, 2090–2101 (2014)

Chapter 10

Ultrafast Nonlinear Optics in the Mid-Infrared



Alexei M. Zheltikov

Abstract The mid-infrared spectral range is unique in many ways. Within this region, electromagnetic radiation can resonate with the most intense signature molecular bands, thus drastically enhancing the coupling between the field and molecular motions. Electrons driven by intense ultrashort mid-IR field waveforms acquire unusually high ponderomotive energies within a fraction of the field cycle, giving rise to new regimes of high-field nonlinear optics. The λ^2 scaling of phase-space mode volume with radiation wavelength λ translates into the λ^2 dependence of the self-focusing threshold, allowing much higher peak powers to be transmitted in a single laser filament in the mid-IR range without losing beam continuity and spatial coherence. Recent breakthroughs in the generation of high-intensity ultrashort pulses in the mid-IR help understand complex interactions of high-intensity ultrashort mid-IR pulses with matter, offer new approaches for coherent and incoherent x-ray generation, enable mid-IR laser filamentation in the atmosphere, facilitate lasing in filaments, give rise to unique regimes of laser-matter interactions, and reveal unexpected properties of materials in the mid-IR range.

A. M. Zheltikov (✉)

Department of Physics and Astronomy, Texas A&M University,
College Station, TX 77843, USA
e-mail: zheltikov@physics.tamu.edu; zheltikov@physics.msu.ru

A. M. Zheltikov

Physics Department, International Laser Center, M. V. Lomonosov
Moscow State University, Moscow 119992, Russia

A. M. Zheltikov

Russian Quantum Center, ul. Novaya 100, Skolkovo, Moscow Region 143025, Russia

A. M. Zheltikov

Kurchatov Institute National Research Center, pl. akad. Kurchatova 1, Moscow 123182,
Russia

10.1 Introduction

Motivated and driven by numerous applications and long-standing challenges in strong-field physics [1, 2], molecular spectroscopy [3, 4], semiconductor electronics [5], and standoff detection [6], ultrafast optical science is rapidly expanding toward longer wavelengths in quest for technologies enabling the generation of high-peak-power ultrashort pulses in the mid-infrared (mid-IR) range. While laser sources of high-power ultrashort pulses in the near-infrared benefit from remarkably efficient laser material with a gain band sufficient for the generation of few-cycle pulses, such materials are lacking in the mid-IR, making it difficult to extend ultrafast laser technologies into the mid-IR range. Within the past two or three years, it has been shown, however, that highly efficient sources of ultrashort pulses in the mid-IR can be created using nonlinear-optical parametric frequency conversion of ultrashort pulses. Optical parametric chirped pulse amplification (OPCPA) has been shown to open new horizons in ultrafast optics in the mid-IR, providing a method whereby sub-100-fs pulses with energies at the level of at least tens of millijoules can be delivered at a central wavelength of about $4\text{ }\mu\text{m}$ as an output of a robust, solid-state compact light source [7]. Such OPCPA-based mid-infrared systems offer unique opportunities for highly efficient coherent [8] and incoherent [9] x-ray generation, open new routes toward subattosecond pulse synthesis [8], enable mid-IR laser filamentation in the atmosphere [10], help achieve lasing in filaments [6] and create high-power light bullets [11], provide a bright source of multi octave high-energy supercontinua [12, 13], and reveal unique regimes of laser-matter interactions [14] along with unexpected optical properties of materials in the mid-IR [15]. A strongly coupled nonlinear spatiotemporal dynamics of ultrashort mid-infrared pulses undergoing self-focusing simultaneously with soliton self-compression in an anomalously dispersive, highly nonlinear solid semiconductor has been shown to enable the generation of multi octave supercontinua with spectra spanning the entire mid-infrared range [16] and compressible to subcycle pulse widths [17]. This chapter offers a brief review of the main achievements and challenges of this rapidly growing field of ultrafast optical science.

10.2 Optical Physics in the Mid-Infrared: Why Longer Wavelengths Make a Difference

The wavelength is one of the key variables in optical science. The properties of materials and regimes of light-matter interactions can vary dramatically from one wavelength range to another. The mid-infrared spectral range is especially difficult in this regard. Ultrafast laser technologies provide a vast arsenal of tools to understand the optical response of matter as a function of the wavelength and offer unique resources to explore the unknown land in the mid-infrared. With the advent

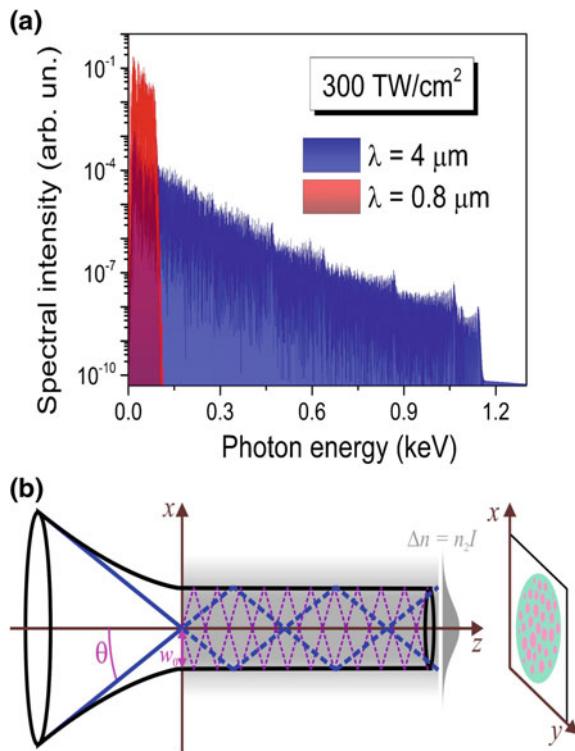
of efficient laser sources of ultrashort pulses in the mid-infrared [8, 18, 19], ultrafast nonlinear optics rapidly expands beyond experiments with a near-infrared driver, revealing unique regimes of optical-harmonic generation [8, 15], laser-induced filamentation [12, 18–20], and high-brightness x-ray generation [9], as well as shedding light on unusual properties of materials in the mid-infrared [14].

Research into the strong-field physics behind light–matter interactions reveals physically significant tendencies and scaling laws in the optical response of matter as a function of the laser driver wavelength λ . We highlight here the wavelength scaling laws of two classes. The scaling laws of the first class deal with electron wave-packet dynamics. In many important cases, a quasi-classical treatment provides an adequate description of this dynamics, leading to a λ^2 scaling for the ponderomotive energy U_p of an electron wave packet, as dictated by the Newtonian mechanics [21, 22]. This scaling reflects fundamental aspects of light–matter interactions and is instrumental in understanding laser electron acceleration, dynamics of recolliding photoelectrons, and high-harmonic generation. The $U_p \propto \lambda^2$ scaling suggests the ways toward a higher spatial resolution in attosecond molecular imaging [23–25] and higher coherent x-ray yields in high-harmonic generation [1, 26] (Fig. 10.1a). Both semiclassical and quantum treatments show that electron wave-packet dynamics within the field half-cycle is the key factor behind the wavelength scaling of optical nonlinearities, explaining the difficulties of the perturbative, phenomenological treatment of ultrafast nonlinear phenomena in terms of nonlinear susceptibilities in the long-wavelength range [14].

The physics behind the scaling laws of the second class is different, as these laws reflect the wave aspects of light–matter interactions and relate to both linear and nonlinear propagation effects. Below in this section, we offer a Hamiltonian-optics perspective on these phenomena, highlighting similarities between the wave equation for optical fields and equations for the quantum-mechanical wave function [27]. As one of the remarkable results of this analysis, the role of the radiation wavelength λ in optics will be shown to be similar, when put in the context of commutation relations and related Heisenberg-type uncertainty, to the role of the Planck constant \hbar in quantum mechanics, as both define the granularity of the phase space defined in terms of the pertinent canonical variables. This λ -dependent phase-space granularity manifests itself in nonlinear-optical propagation effects, such as self-focusing, where the spatial modes of light show up through a breakup of a coherent beam into N filaments (Fig. 10.1b), with N controlled by the number of self-trapped spatial modes within the beam. This perspective sheds light on the self-focusing threshold P_s as the peak power needed to trap a single spatial mode of an optical field, with the λ^2 dependence of the phase-space volume of this mode translating into the celebrated λ^2 scaling of P_s .

To illustrate how the wavelength λ limits the elementary phase-space volume in optics, we consider a Gaussian beam of radiation with a beam radius w_0 . Diffraction-induced divergence of this beam is characterized by a divergence angle $\theta \approx \lambda(\pi w_0)^{-1}$. The phase-space mode volume, Ξ_m , defined as a product of the beam area $S_m = \pi w_0^2$ and the solid angle of beam divergence, $\Omega_m = \pi\theta^2$, is then a constant,

Fig. 10.1 **a** Spectra of high-order harmonics calculated by numerically solving the time-dependent Schrödinger equation for a helium atomic target irradiated by a laser pulse with a field intensity of 300 TW/cm^2 and a central wavelength of $0.8 \mu\text{m}$ (red) and $4.0 \mu\text{m}$ (blue). **b** Self-focusing couples a freely diffracting laser beam to a group of spatial modes (solid and dashed lines) trapped by the refractive index change $\Delta n = n_2 I$. The interference of these modes, partially randomized by stochastic processes involved in beam dynamics, including ionization, gives rise to a speckle pattern



$$\Xi_m = S_m \Omega_m = \lambda^2, \quad (1)$$

controlled by the radiation wavelength.

With time t in (1) replaced by the longitudinal coordinate z , the Fermat's principle in optics becomes mathematically identical to the Hamilton's principle in mechanics [28, 29]. Optical rays can then be treated as trajectories of classical particles, while the evolution of optical waves has much in common with the dynamics of quantum-mechanical wave functions. With the Hamiltonian formulation of the Fermat's principle,

$$\int_{z_1}^{z_2} L(x, y, x', y', z) dz = \min, \quad (2)$$

where

$$L(x, y, x', y', z) dz = n(x, y, z) [1 + x'^2 + y'^2]^{1/2} \quad (3)$$

is the Lagrangian, $x' = dx/dz$, $y = dy/dz$, and $n(x, y, z)$ is the refractive index, the generalized momenta conjugate to the variables x and y are given by [30]

$$p_x = \frac{\partial L}{\partial x'} = n \sin \alpha_x, \quad (4)$$

$$p_y = \frac{\partial L}{\partial y} = n \sin \alpha_y, \quad (5)$$

with α_x and α_y being the angles defining the ray direction.

The Liouville theorem applied to a bundle of rays requires the phase-space volume filled with points representing this bundle to remain constant. We can now relate the constancy of the phase-space mode volume Ξ_m in (1) to the Liouville theorem applied to the x, y, p_x, p_y phase space. The fundamental significance of this argument is, perhaps, best appreciated in terms of the second law of thermodynamics, which prohibits any linear optical system to change the phase-space volume of a spatial mode. Furthermore, with momenta p_x and p_y defined as $\hat{p}_x = -i\lambda \frac{\partial}{\partial x}$, $\hat{p}_y = -i\lambda \frac{\partial}{\partial y}$, which offers a quantum version of (4) and (5), the canonically conjugate variables x, y, p_x , and p_y satisfy commutation relations $xp_x - p_x x = i\lambda$ and $yp_y - p_y y = i\lambda$, which translate into the Heisenberg-type uncertainty, $(w_0\theta)^2 \geq (\lambda/\pi)^2$, in perfect agreement with (1).

When written in a canonical form,

$$ik \frac{\partial E}{\partial z} + \frac{1}{2} \Delta_{\perp} E + \eta |E|^2 E = 0, \quad (6)$$

the nonlinear Schrödinger equation (NSE), which is applicable to both a quantum system in a nonlinear potential and an electric field in a nonlinear medium, reveals two physically significant integrals of motion [31]—the energy integral,

$$Q = \frac{1}{2k^2} \int_{-\infty}^{\infty} \left[(\Delta_{\perp} E)^2 - \eta |E|^4 \right] d^2\rho, \quad (7)$$

and the integral of the number of particles,

$$N = \int_{-\infty}^{\infty} |E|^2 d^2\rho, \quad (8)$$

where ρ is the radius vector in the xy -plane.

With the physical units restored, the N integral of (8) can be represented as $N = P/P_s$, where $P_s = \lambda^2 (2\pi n_0 n_2)^{-1}$ is the critical power of self-focusing [31–33]. The

N integral can thus be understood as the number of modes trapped by the refractive index change $\Delta n = n_2 I$ induced by a laser beam with an on-axis field intensity I [34]. When the peak power of a laser pulse is exactly equal to the critical power of self-focusing, $P = P_s$, we find that $N = 1$. This implies that the refractive index change induced by a laser beam with $P = P_s$ is sufficient to confine only one spatial mode. The λ^2 scaling of the self-focusing threshold P_s can now be understood as a reflection of the $\Xi_m \propto \lambda^2$ scaling of the phase-space volume of spatial modes.

10.3 Ultrafast Optical Nonlinearities in the Mid-Infrared: The Role of Ionization Dynamics Within the Field Half-Cycle

The main goal of this section is to shed light on unusual properties of optical nonlinearities in the mid-IR. To this end, we analyze the key tendencies in the behavior of optical nonlinearity of a generic hydrogen-like quantum system as a function of the wavelength of the driver field. This analysis shows that, as the wavelength of the driver is increased, free-state electrons start to dominate over bound electrons in the overall nonlinear response. In this regime, the main properties of the nonlinear response of a quantum system can be adequately understood in terms of a semiclassical nonlinear electron current modulated by the driver field [14].

For the purposes of our study, it is convenient to represent the general solution to the time-dependent Schrödinger equation (TDSE) for an electron in a hydrogen-like atom in the presence of an external electric field as a sum $\psi(\vec{r}, t) = \psi_b(\vec{r}, t) + \psi_f(\vec{r}, t)$ of positive- and negative-energy terms, $\psi_b(\vec{r}, t)$ and $\psi_f(\vec{r}, t)$, corresponding to the bound and free (continuum) states of an electron. The bound-state part of $\psi(\vec{r}, t)$ can be expanded as $\psi_b(\vec{r}, t) = \sum_{n=1}^N \sum_{l=0}^{n-1} \alpha_{n,l}(t) \psi_{n,l}(\vec{r})$, where $\psi_{n,l}(\vec{r})$ are the orthonormalized eigenfunctions of the pertinent stationary Schrödinger equation, with the probability to find an electron in a bound state with quantum numbers n and l given by $|\alpha_{n,l}(t)|^2$. The probability to find an electron in the continuum, or population of the continuum, at the instant of time t can be found as $\rho_f(t) = \int_V |\psi_f(\vec{r}, t)|^2 d\vec{r}$.

The z -component of the dipole moment, $d_z(t) = \int_V e \psi^*(\vec{r}, t) z \cos(\theta) \psi(\vec{r}, t) d\vec{r}$, is then given by a sum of three terms, $d_z(t) = d_{bb}(t) + d_{ff}(t) + d_{bf}(t)$, which isolate the contributions of bound-bound, free-free, and bound-free electron transitions, respectively. The separation of the dipole moment into three parts is much more than a matter of convenience. These components will be shown to have distinctly different underlying physics and will help reveal regimes where important properties of the nonlinear-optical response of a quantum system can be understood in terms of classical models.

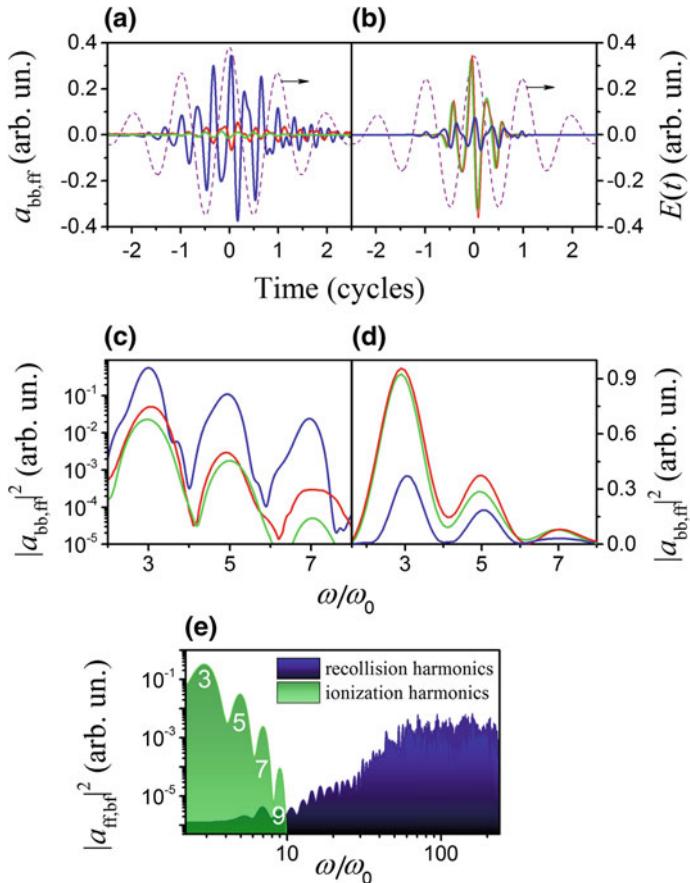


Fig. 10.2 Time-resolved radiation amplitudes a_{bb} (blue) and a_{ff} (red) of bound-state and free electrons (**a**, **b**) and their spectra (**c**, **d**) for a laser driver (purple dashed line) with the field intensity $I_0 = 110 \text{ TW/cm}^2$, the pulse width $\tau_0 = 10/\omega_0$, and the central wavelength $\lambda_0 = 0.8 \mu\text{m}$ (**a**, **c**) and $4.0 \mu\text{m}$ (**b**, **d**). Results of calculations using the semiclassical model are given with a green line. **e** The spectra of a_{ff} (green) and a_{bf} (blue and purple) calculated using the TDSE for a laser driver pulse with $I_0 = 110 \text{ TW/cm}^2$, $\tau_0 = 10/\omega_0$, and $\lambda_0 = 0.8 \mu\text{m}$ (purple) and $4.0 \mu\text{m}$ (blue and green). Reproduced from [14]

In Fig. 10.2, we present typical time dynamics (Fig. 10.2a, b) and spectra (Fig. 10.2c–e) of radiation amplitudes $a_{bb, ff, bf} \propto \partial^2 d_{bb, ff, bf} / \partial t^2$ calculated for different intensities, I_0 , and carrier wavelengths, λ_0 , of the driver field. The properties of radiation spectra emitted by different parts of the dipole moment are strikingly different, providing a map of significant differences in their physical nature. The intensity of low-order harmonics in the spectrum of a_{bb} (blue line in Fig. 10.2c, d) rapidly decays as a function of the harmonic order, indicating that, even for relatively intense laser fields, the driver field can still be treated perturbatively for

bound-state electrons, which always remain near the atomic core, with the energy these electrons gain from an external driver being always less than the energy of their interaction with the core.

The harmonic spectrum of a_{bf} sends an opposite message, as it displays an extended plateau (Fig. 10.2e), which can span over hundreds of harmonics, followed by a cutoff. These features of harmonic spectra, widely exploited in attosecond technologies [34], clearly indicate the nonperturbative character of nonlinear-optical effects accompanying the electron rescattering process [35], as discussed in the extensive literature following the seminal work by Corkum [21]. The shape of the spectrum of this part of nonlinear-optical response is controlled by the ionization potential and the ponderomotive energy acquired by electrons before the recollision. The spectra of low-order harmonics due to the a_{ff} term follow a radically different scaling. For weak fields and short carrier wavelengths, the a_{ff} term is always smaller than a_{bb} (Fig. 10.2c). However, as I_0 and/or λ_0 are increased, free electrons tend to play a progressively more important role (Fig. 10.2d), until, eventually, in the regime of high intensities and long carrier wavelengths, these electrons start to dominate low-order harmonic generation (Fig. 10.2d).

To understand these tendencies, it is instructive to examine the buildup of the density of free electrons within the field half-cycle, controlled by population of the continuum $\rho_f(t)$. For low field intensities and short driver wavelengths, the electron wave function is strongly localized around the core. In this regime, the electrons spend much time near the core, translating into a strongly oscillatory behavior of the continuum population, which shows that most of the electrons that undergo ionization within the field half-cycle tend to recombine, refilling the bound states after this field half-cycle and keeping the continuum population low. The harmonic spectra in Fig. 10.2c are very instructive as they clearly indicate that the nonlinear response in this regime is dominated by bound-state electrons.

For higher field intensities and/or longer central wavelengths of the driver, electrons travel further away from the core, acquiring a higher energy within each field half-cycle. This tends to delocalize the electron wave function. A significant fraction of electrons undergoing ionization does not recombine, giving rise to a stepwise buildup of continuum population after each field half-cycle with much higher free-electron densities in the wake of the driver pulse. The steps in the continuum population synchronized with field half-cycle translate into odd harmonics of the driver in the spectral domain. The intensity of these harmonics is seen to be much higher than the intensity of harmonics due to bound-state electrons (Fig. 10.2b, d).

This analysis shows [14] that subcycle ionization dynamics, or, more specifically, the dynamics of continuum population within the field half-cycle, is the key physical factor that controls the properties of optical nonlinearity as a function of the carrier wavelength and intensity of a driving laser field. For high-intensity low-frequency fields, free-state electrons dominate over bound electrons in the overall nonlinear response of a quantum system. In this regime, the main properties of the nonlinear response of a quantum system can be adequately understood in terms of a semiclassical nonlinear electron current modulated by the driver field.

While transitions between free and bound electronic states enable high-order harmonic generation and attosecond technologies [36], nonrecolliding tunneling free-state electron wave packets driven by a low-frequency field can help synthesize, as shown in [37], extremely short field waveforms lasting a tiny fraction of the field cycle with intensities orders of magnitude higher than the intensities of high-harmonic pulses. A low-frequency, mid-infrared ultrashort intense driver field can give rise to high-contrast isolated tunneling electron wave packets, capable of radiating subcycle field waveforms as a part of low-order nonlinear-optical processes. In a striking contrast with nonlinear optics in the near-infrared, the subcycle dynamics of electron wave packets induced by mid-infrared drivers can dominate over the bound-electron contribution to low-order harmonic generation, enabling the generation of subcycle field waveforms whose intensity is orders of magnitude higher than the intensity of high-harmonic pulses.

The sensitivity of the field waveform of low-order harmonics emitted by free-state electrons to the shape of the cycle of the driver field suggests cycle engineering as a powerful tool for a synthesis of extremely short, high-contrast optical pulses. To demonstrate this method of short-pulse synthesis, we use a tailored driver field consisting of a laser pulse with a carrier wavelength $\lambda_0 = 4 \mu\text{m}$, the field amplitude E_1 , and a pulse width of 27 fs mixed with a second harmonic pulse with a central wavelength of $\lambda_0/2 = 2 \mu\text{m}$, the field amplitude E_2 , and a phase shift $\Delta\phi$ relative to the fundamental field. A short and intense central peak of such a tailored field waveform produced with $E_2/E_1 = 0.13$ and $\Delta\phi = 0$ gives rise to a steep ionization step (green line in Fig. 10.3a), inducing an intense isolated tunneling electron wave packet in the continuum. Such an abrupt ionization process leads to a radiation of a broadband, supercontinuum spectrum without any well-resolved harmonic features. With an appropriate spectral filter, this radiation spectrum yields a field waveform whose pulse width is 10 times shorter than the field cycle (Fig. 10.3d). This central pulse is preceded by a lower amplitude spike whose peak intensity is more than five times lower than the peak intensity of the central peak. The entire field waveform produced as a part of this process, as well as its pulse width are efficiently controlled by the E_2/E_1 ratio (Fig. 10.3a–f) and the phase shift $\Delta\phi$ of the second harmonic relative to the fundamental field.

With an appropriate compression of individual pulses, the coherent combining of these fields yields cycle-tailored field waveforms. With appropriate optimization of the amplitudes and phases of the constituent fields, as well as the delay time between the pulses, only two central peaks of this field waveform are powerful enough to drive high-order harmonic generation, with the first cycle being much shorter than the second cycle (Fig. 10.4). Now, the first, shorter peak can efficiently ionize an atom, while the second, longer peak is ideally suited to accelerate the electron wave packet and drive it back to the parent ion, giving rise to very-high-ponderomotive-energy recolliding electrons. The field tailoring process can be steered, e.g., by using a blind search with genetic algorithms, toward high-power attosecond pulse generation or keV x-ray supercontinua.

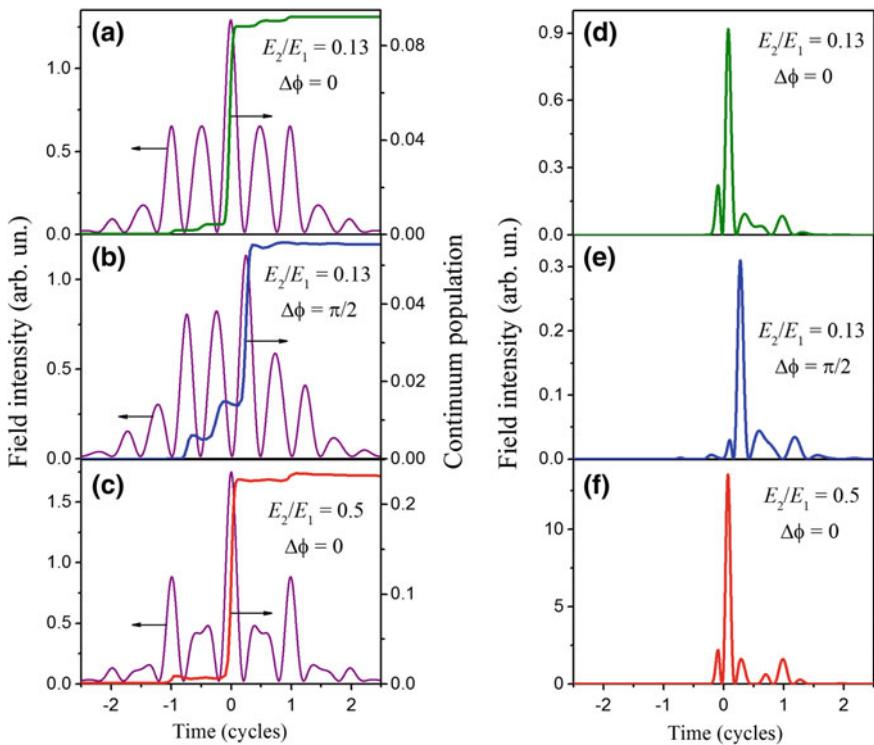


Fig. 10.3 **a, b, c** The driver field (green) and the continuum population (blue) as a function of time and **d, e, f** the field behind the spectral filter for $E_2/E_1 = 0.13$ (**a, b, d, e**) and 0.5 (**c, f**) and $\Delta\phi = 0$ (**a, c, d, f**) and $\pi/2$ (**b, e**). Reproduced from [37]

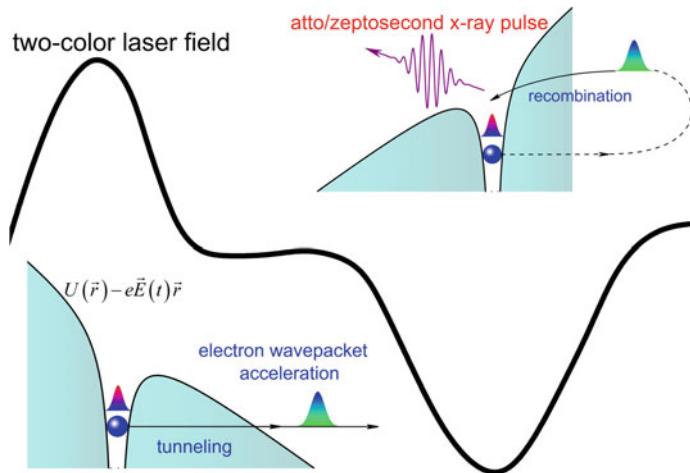


Fig. 10.4 Generation of ultrashort attosecond pulses with a cycle-tailored field waveform

10.4 Mid-Infrared Laser Filaments in the Atmosphere

Laser-induced filamentation [38–40] is a thrilling phenomenon of ultrafast optical physics, in which diffraction of a laser beam is suppressed by a combined effect of self-focusing and transverse electron density profile induced by ultrafast photoionization. While filamentation of ultrashort light pulses with peak powers above the self-focusing threshold is a universal phenomenon, observed in gases, liquids, and solids [39, 40], laser filaments in the atmosphere are of special significance as they offer unique opportunities for long-range signal transmission, delivery of high-power laser beams, and remote sensing of the atmosphere [41].

In Fig. 10.5, we present a survey of representative laser filamentation experiments [12, 20, 38–40, 42–50] in various gases, including the atmospheric air, in a diagram where the wavelength and the peak power of laser pulses are used as coordinates. The single-filamentation regime in the atmospheric air exists in the dark area of this diagram. Loosely defined boundaries of this area are shown by gradient shading. The lower boundary is centered at the critical power of self-focusing [32, 51], $P_{\text{cr}} = C(8\pi n_0 n_2)^{-1} \lambda^2$, where n_0 is the field-free refractive index, n_2 is the nonlinear refractive index and C is a numerical factor, $3.72 < C < 6.4$, defined [52] by the beam profile ($C \approx 3.72$ and 3.77 for Townesian and Gaussian beams, respectively). The P_{cr} threshold is the key parameter for laser-induced filamentation, which sets a fundamental limit on the peak power and, hence, the energy of laser pulses in a filament. Below this threshold (purple shaded area in Fig. 10.5), the nonlinear lens induced by a laser beam is not strong enough to compensate for beam diffraction. In the opposite case, when the peak power of a laser field becomes much higher than P_{cr} (rose shading in Fig. 10.5), the beam tends to decay into multiple small-scale filaments [39, 40, 53, 54], eventually losing its spatial coherence. The parameter space where a laser beam can propagate as a single filament in the atmospheric air is thus bound to the dark shaded area in the diagram of Fig. 10.5.

Until recently, experiments on laser-induced filamentation in the atmospheric air (white circles in Fig. 10.5) were limited to the visible and near-infrared ranges ($\lambda < 1030$ nm), where sufficiently powerful short-pulse laser sources were available [38–47]. Most of those earlier experiments on laser filamentation in the atmosphere were performed using Ti: sapphire laser systems [38–44]. This class of lasers can deliver ultrashort pulses within a broad range of peak powers, allowing the generation of single and multiple filaments in air (white and rose segments of an ellipse centered at 800 nm in Fig. 10.5), as well as a variety of filamentation regimes in high-pressure gases (the yellow segment of the ellipse at 800 nm). Because of the λ^2 scaling of the critical power of self-focusing P_{cr} , using a longer-wavelength laser driver is a straightforward strategy for increasing the laser peak power and radiation energy in a single filament. Overall, filamentation experiments in high-pressure atomic and molecular gases performed using laser drivers with different carrier wavelengths [12, 20, 38–40, 42–50] confirm this possibility (Fig. 10.5). However, building longer- λ alternatives to Ti: sapphire [38–44] and, since recently, ytterbium

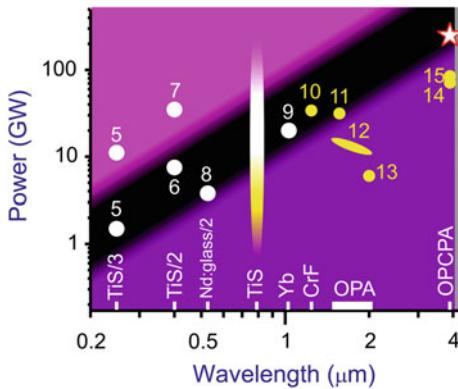


Fig. 10.5 Survey of filamentation experiments in gas media [12, 20, 38–40, 42–50]. The peak power of laser pulses used for laser-induced filamentation is shown versus the central wavelength of these pulses for experiments performed in the atmospheric air (white circles) and in high-pressure gases (yellow circles): high-pressure N₂ [47], Ar at a pressure of 5 bar [48], Kr at 4 bar [49], Xe at 2.1 bar [50], Ar at 4.5 bar [12], O₂ and N₂ at bar [20]. References are given by numbers. Single-filamentation regime in the atmospheric air exists in the dark area of the diagram. Its loosely defined boundaries are shown by gradient shading, with the lower boundary centered at the critical power of self-focusing, $P_{\text{cr}} = C(8\pi n_0 n_2)^{-1} \lambda^2$, $C \approx 6.4$, and the upper boundary centered at $7P_{\text{cr}}$. Above the upper boundary, a laser beam tends to break up into multiple filaments. Continuum of filamentation experiments using Ti: sapphire (TiS) laser systems [38–40] is shown by the ellipse centered at 800 nm. Other sources of ultrashort pulses for filamentation experiments (shown along the abscissa axis) include Yb: CaF₂ laser (Yb), Cr: forsterite laser (CrF), optical parametric amplifier (OPA), optical parametric chirped-pulse amplifier (OPCPA), the second harmonic of a Nd: glass laser (Nd: glass/2), and the second (TiS/2) and third (TiS/3) harmonics of a Ti: sapphire laser output. Filamentation experiments presented in this work are shown by a star. The grey area on the right represents the atmospheric CO₂ absorption band. Reproduced from [18]

[46] and Cr: forsterite [47] lasers that would be capable of delivering ultrashort laser pulses with peak powers above P_{cr} for the atmospheric air, i.e., at least a factor of λ^2 higher than the peak powers of amplified Ti: sapphire and ytterbium laser pulses used for atmospheric filamentation, is a challenging problem.

Cutting-edge laser technologies, based on optical parametric chirped-pulse amplification (OPCPA) in the mid-infrared, offer powerful tools to confront this challenge. In experiments presented in [10], a high-peak-power compressed output of a multistage mid-infrared OPCPA system (Fig. 10.6) is employed to enable the generation of mid-infrared laser filaments in the atmosphere for the first time. The stretched-pulse OPCPA output in these experiments had an energy above 50 mJ. Compression of these pulses using a grating compressor yielded mid-infrared pulses with a pulse width of 80–200 fs and an energy up to 30 mJ. The 3.9-μm OPCPA output with a peak power exceeding $P_{\text{cr}} \approx 150$ GW loosely focused by a CaF₂ lens with a focal length ranging from 0.5 to 1.2 m (Fig. 10.6a) induces a filament in the atmosphere. Filament formation is visualized by a bright spark whose length varies from a few centimeters up to 30 cm, depending on the peak power of the laser driver and the focal length of the focusing

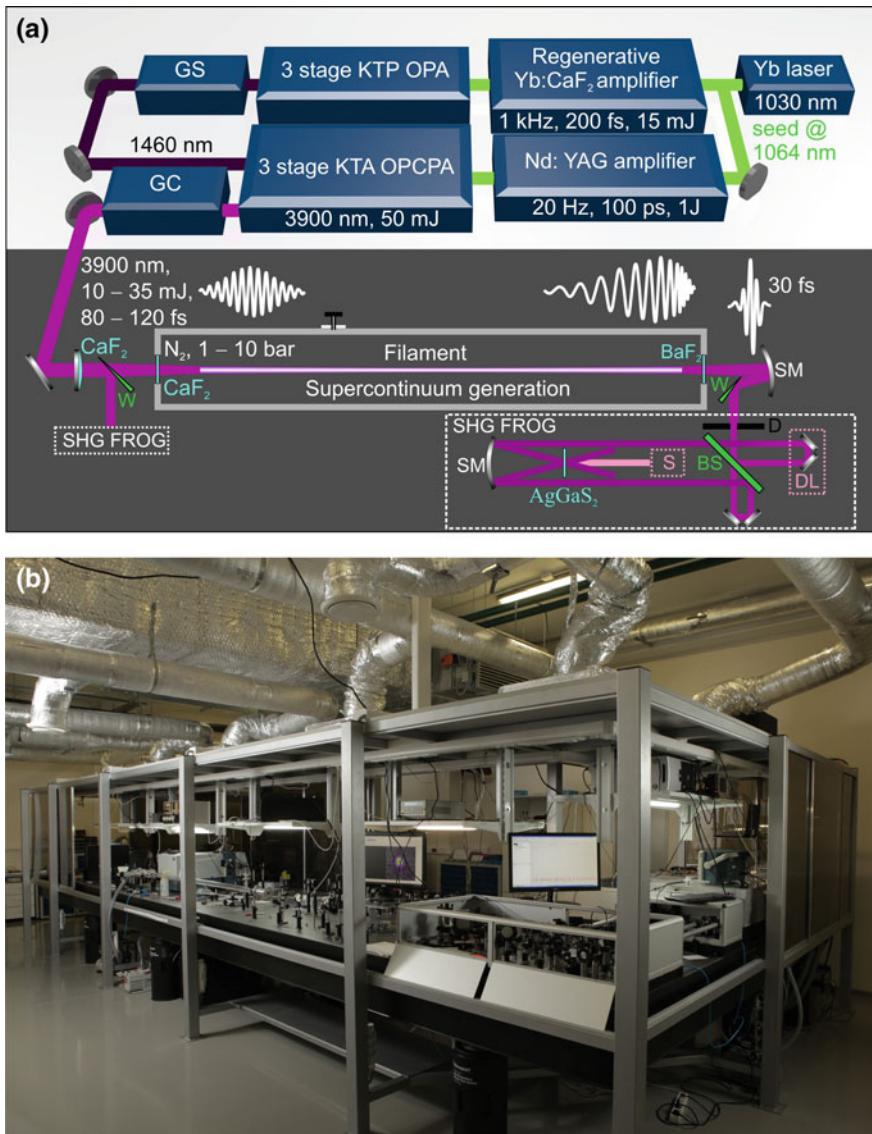


Fig. 10.6 **a** Filamentation and supercontinuum generation with high-power mid-IR pulses: GS, grism stretcher; GC, grating compressor; MIR spec, mid-IR spectrometer, D, diaphragm; SM, spherical mirror; W, CaF₂ wedge, BS, thin-film beam splitter; DL, tunable delay line; S, spectrometer; PD, photodetector. **b** The laser facility with an OPCPA source of subterawatt ultrashort pulses in the mid-infrared at the Advanced Photonics Laboratory of the Russian Quantum Center

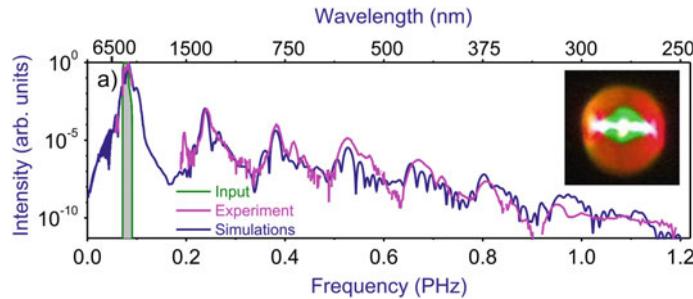


Fig. 10.7 The on-axis spectrum of supercontinuum radiation generated in a filament induced by the mid-IR OPCPA output in the air: (red) experiment, (blue) simulations. The spectrum of the mid-IR driver inducing the filament is shown with grey shading. The supercontinuum beam behind the filament is shown in the inset. Reproduced from [13]

lens. Filamentation of laser pulses is accompanied by a dramatic increase in the bandwidth of the mid-infrared pulse behind the region of filamentation. A typical spectrum of such broadband radiation measured in our filamentation experiment in the atmosphere is shown in Fig. 10.7. The long-wavelength part of the supercontinuum spectra (Figs. 10.7 and 10.8a) represents a broadened spectrum of the mid-infrared driver with an extended blue wing and a long-wavelength tail attenuated and eventually limited by the absorption of carbon dioxide in the atmosphere. The near-infrared section of the spectrum (Fig. 10.8a–c) is dominated by spectrally broadened peaks observed near the frequencies of the third and fifth harmonics of the mid-infrared driver. Finally, the visible part of the spectrum (Figs. 10.7 and 10.8a–c) features prominent maxima near the frequencies of the seventh and ninth harmonics of the driving field. A typical far-field beam profile of the filament output, shown in the insets to Figs. 10.7 and 10.8a, reveals a bright white-light central part surrounded by greenish and reddish outer rings.

To analyze the spatiotemporal field dynamics leading to a filamentation of ultrashort mid-infrared pulses, we use a model based on the nonlinear field evolution equation [39, 40, 55, 56] that includes the dispersion of the medium, beam diffraction, optical nonlinearities due to the third-, fifth-, seventh-, and ninth-order susceptibilities of a gas medium, ionization-induced nonlinearities, pulse self-steepening, spatial self-action phenomena, as well as plasma-related loss, refraction, and dispersion. This equation is solved jointly with the equation for the electron density $\rho(t)$ with the photoionization rate calculated using the Popov-Perelomov-Terentyev version of the Keldysh formalism. Numerical simulations accurately reproduce the key features in the experimental spectra (Figs. 10.7 and 10.8a–c) and beam profiles (Fig. 10.8d) of the filament output and provide an accurate estimate for the filament length (Fig. 10.8d). This numerical analysis reveals a number of striking features in the nonlinear spatiotemporal dynamics of mid-infrared pulses in the filamentation regime that never show up in near-infrared laser filaments.

As can be seen from Fig. 10.6c, the initial stage of filamentation of a mid-infrared beam, within the range of the propagation coordinate z from 0 to 60 cm, is dominated by self-phase modulation—a universal mechanism of spectral broadening of ultrashort laser pulses regardless of the spectral range. However, already at this initial stage of filamentation, other significant nonlinear processes come into play, most notably, odd-order optical harmonic generation. While third-harmonic generation is not uncommon to laser filamentation of Ti: sapphire laser pulses in the atmosphere [39, 40, 57], higher order harmonics of a 800-nm driver fall outside the atmospheric transmission window and do not show up in filamentation in the near-infrared. The situation is drastically different in the case of filaments induced by mid-IR pulses, when a whole group of odd-order harmonics can fall, as in the case of our experiments, within the transmission range of the atmospheric air. Moreover, the weakness of dispersion of the atmospheric air in the mid-infrared favors phase matching for the generation of these harmonics. These two factors give rise to a unique scenario of optical harmonic generation in a mid-infrared laser filament. The near-infrared and visible parts of filament output spectra in our experiments feature prominent signals near the frequencies of the third, fifth, seventh, and ninth harmonics of the 3.9- μm driver, all falling within the atmospheric transparency range. Combined with the mid-infrared supercontinuum, these harmonics give rise to remarkably broad radiation spectra at the output of the filament, spanning the entire visible and near-infrared ranges and covering a considerable part of the mid-infrared up to the edge of the mid-infrared atmospheric transmission window.

These insights into the beam dynamics help understand the ring structure of the far-field beam profile observed in our mid-infrared filamentation experiments. In the central part of the beam (direction labeled with arrow 1 in Fig. 10.8d), generation of a mid-infrared supercontinuum is accompanied by efficient generation of the third, fifth, seventh, and ninth harmonics in a collinear geometry, giving rise to a bright white spot centered on the beam axis at the center of the far-field beam pattern (the insets in Figs. 10.7 and 10.8a). Since the third harmonic lies in the near-infrared range, it does not contribute to the colors of the rings in the output beam pattern. Within the range of angles 1.3–3.5 mrad (direction 2 in Fig. 10.8d), the fifth harmonic also falls outside the visible range, with the seventh harmonic, centered at 0.56 μm (Fig. 10.8b), giving rise to a greenish color of the beam pattern. Finally, at the periphery of the beam (corresponding to direction 3 in Fig. 10.8d), the intense fifth-harmonic signal, blue-shifted to 0.65 μm (Fig. 10.8c), is responsible for a reddish color of the outer ring in the output beam profile observed in experiments (the insets in Figs. 10.7 and 10.8a). The colors of the far-field beam pattern synthesized from the results of numerical simulations (Fig. 10.8d, right) are fully consistent with experimental beam profiles (the insets in Figs. 10.7 and 10.8a).

To summarize, filamentation of ultrashort mid-infrared pulses in the atmosphere has been demonstrated for the first time. With the spectrum of a femtosecond laser driver centered at 3.9 μm , right at the edge of the atmospheric transmission window, radiation energies above 20 mJ and peak powers in excess of 200 GW have been transmitted through the atmosphere in a single filament. Experimental studies

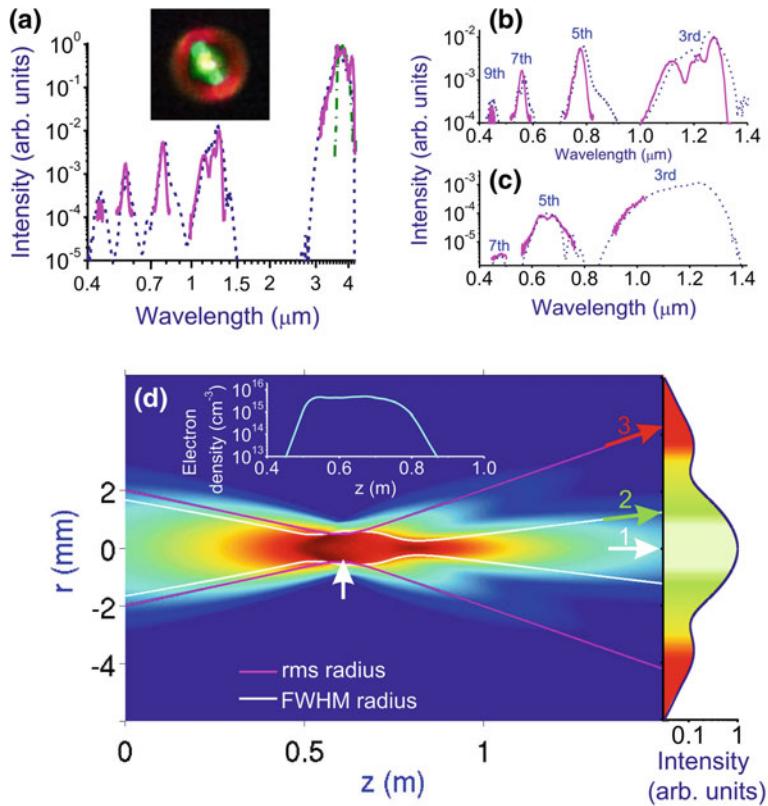


Fig. 10.8 The overall spectrum of radiation behind the filament induced by the mid-infrared driver (a) and its visible-near-infrared part on the beam axis (b) and on the periphery of the beam (c): experiments (red solid line) and simulations (dotted lines). The input spectrum of the mid-infrared driver is shown by the green dash-dotted line. The experimental beam profile is shown in the inset. **d** Beam dynamics of the mid-infrared driver in the filamentation regime. The color-coded field intensity in the mid-infrared driver is calculated as a function of the radial coordinate r and the propagation path z . The focus of the nonlinear lens calculated using the Marburger formula ($z = 60$ cm) is shown with a vertical arrow. The behavior of the full width at half-maximum (FWHM) beam radius (white line) visualizes a formation of a filament in the central part of the beam. The rms beam radius (red line) visualizes enhanced divergence of the outer part of the beam due to the scattering off the electron-density profile induced by the mid-infrared driver. Arrows 1, 2, and 3 indicate directions in which the central bright white spot, the inner green ring, and the outer red ring are observed in the far-field beam pattern, as shown in the plot on the right of the map. The electron density on the beam axis as a function of the propagation path along the filament is shown in the plot in the upper part of the map. Reproduced from [18]

reveal unique properties of mid-infrared filaments, where the generation of powerful mid-infrared supercontinuum is accompanied by unusual scenarios of optical harmonic generation, giving rise to remarkably broad radiation spectra, stretching from the visible to the mid-infrared.

10.5 Mid-Infrared-to-Mid-Ultraviolet Supercontinua

Broadband sources of coherent radiation in the mid-infrared (mid-IR) range are in great demand for a broad variety of applications, including molecular vibrational spectroscopy [3], semiconductor optoelectronics [5], frequency-comb technologies [4], trace-gas detection [58], biomedical diagnostics [59], and food quality control [60]. While they cannot compete with synchrotron-based mid-IR beamlines in terms of brightness [61, 62], rapidly progressing mid-IR laser sources open the ways toward lower costs and real-life-friendly formats. As a particularly promising direction, nonlinear-optical methods of mid-IR supercontinuum generation have been demonstrated using properly chosen solid-state materials [63, 64] and non-silica fibers [65].

Supercontinua with peak powers well above the limits dictated by laser damage thresholds of solid materials can be generated by high-power ultrashort light pulses inducing laser filaments in gas media [39, 40]. These experiments show that laser filaments induced in the air by 0.25-TW, sub-100-fs mid-IR pulses can enable the generation of high-energy multi octave supercontinua (Fig. 10.7). Optical harmonics up to the 15th order are produced in such laser filaments, radically enhancing the high-frequency part of supercontinuum spectra.

10.6 Multi octave Sub-Two-Cycle Supercontinua in the Mid-Infrared from Self-compressing, Self-focusing Soliton Transients in a Solid

In view of the impressive progress of ultrafast optical science and technologies in the mid-infrared, the key question to be addressed is to what extent the nonlinear-optical strategies of pulse compression and coherent broadband waveform generation can be extended to the mid-IR range. In the near-infrared, the physical scenarios behind these processes have been thoroughly understood [66], enabling highly efficient technologies for few- and even single-cycle pulse generation. The extension of these scenarios to the mid-IR is, however, in no way trivial, facing challenges with a complex wavelength scaling of ultrafast nonlinear-optical phenomena and optical nonlinearities of materials. Fiber-optic strategies, which play the key role in ultrafast nonlinear optics in the near-infrared [67], are not easily transferable to the entire mid-IR range because of a prohibitively high loss of fiber materials in this spectral region. As a promising alternative, nonlinear-optical transformation of ultrashort mid-IR pulses in bulk solids attracts much interest [63, 64, 68–72] offers attractive solutions for pulse compression and supercontinuum generation of mid-IR pulses with moderate peak powers.

Experiments reported in [16] employ a frequency-tunable source of ultrashort pulses in the mid-IR, which involves two sequential stages of nonlinear-optical down conversion (Fig. 10.9). At the first stage, 65-fs, 0.8-mJ, 810-nm, 1-kHz

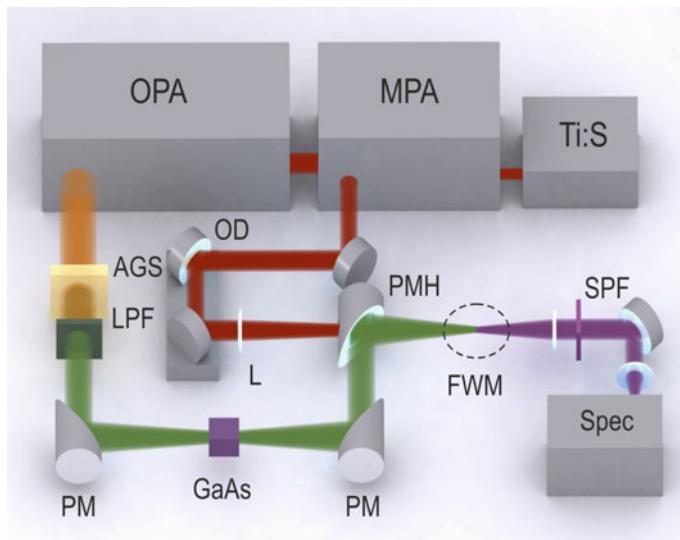


Fig. 10.9 Experimental setup: Ti: S, mode-locked Ti: sapphire master oscillator; MPA, multipass amplifier; OPA, optical parametric amplifier; AGS, AgGaS₂ crystal; LPF, longpass filter; PM, parabolic mirror; L, BK7 glass lens; PMH, parabolic mirror with a hole; FWM, four-wave mixing in a gas medium; SPF, shortpass filter; Spec, spectrometer

pulses delivered by a Ti: sapphire laser, consisting of a master oscillator and a multipass amplifier, are used to produce a broadband seed signal through supercontinuum generation in a sapphire plate and serve as a pump for an optical parametric amplification (OPA) of the seed signal in a BBO crystal, yielding tunable signal and idler. At the second stage, the signal and idler fields generate a mid-IR field wavelength-tunable from 2.85 to 13 μm (Fig. 10.9) through difference-frequency generation (DFG) in an AgGaS₂ (AGS) crystal [73]. Pulse characterization of the mid-IR output is performed via cross-correlation frequency-resolved optical gating (XFROG) based on four-wave mixing (FWM) in a gas medium [64, 74]. To this end, an ultrashort mid-IR pulse is combined with a reference Ti: sapphire laser pulse on an off-axis 100-mm-focal-length parabolic mirror with a hole, which focuses both pulses into a molecular or atomic gas (Fig. 10.9) to generate an FWM signal through the $\omega_{\text{FWM}} = 2\omega_p - \omega_d$ process.

Soliton self-compression of the wavelength-tunable mid-IR DFG output is implemented using GaAs (Fig. 10.9), which possesses a high nonlinearity, $n_2 \approx 3 \cdot 10^{-14} \text{ cm}^2/\text{W}$, and an anomalous dispersion for wavelengths longer than its zero group-velocity-dispersion wavelength, $\lambda_z \approx 6.8 \mu\text{m}$. Nonlinear transformation of mid-IR pulses in GaAs gives rise to supercontinuum radiation, accompanied by pulse self-compression with pulse compression ratios exceeding 3 (Fig. 10.10c, f). Specifically, a mid-IR DFG output with $\tau_0 \approx 150 \text{ fs}$, $\lambda_0 \approx 7.9 \mu\text{m}$, and $W_0 \approx 2 \mu\text{J}$ undergoes spectral broadening and self-compression in a 5-mm GaAs plate,

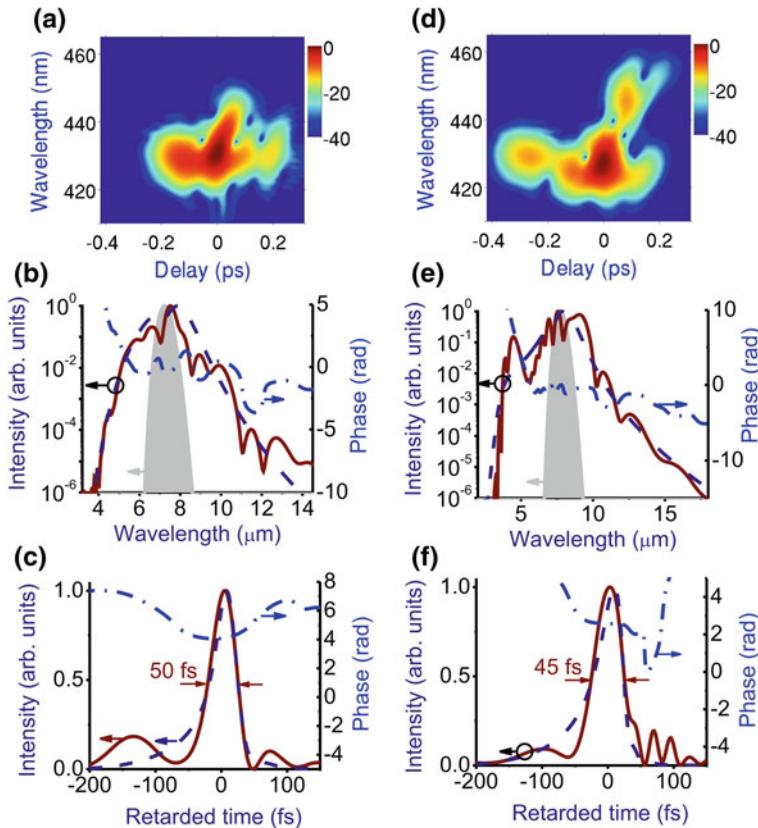


Fig. 10.10 Self-compression of mid-IR pulses with an initial pulse width $\tau_0 \approx 150$ fs, central wavelength $\lambda_0 \approx 7.1$ μm (**a–c**) and 7.9 μm (**d–f**), and input energy $W_0 \approx 2$ μJ in a 5-mm GaAs plate: **a**, **d** an FWM XFROG trace of the mid-IR pulse behind the GaAs plate; **b**, **e** the input spectrum of the pulse (grey shading), experimental (solid line) and simulated (dashed line) output spectra, and the spectral phase (dash-dotted line) of the output waveform retrieved from the FWM XFROG trace; **c**, **f** experimental (solid line) and simulated (dashed line) temporal envelope of the mid-IR pulse behind the GaAs plate with the phase (dash-dotted line) retrieved from the FWM XFROG trace. Reproduced from [16]

evolving to a waveform with a pulse width $\tau_c \approx 45$ fs behind this plate (Fig. 10.10f). Such a pulse contains only 1.2 cycles of the field at $\lambda_0 \approx 7.9$ μm .

To understand the spatiotemporal dynamics of ultrashort pulses behind the generation of sub-two-cycle field waveforms in the mid-IR in a highly nonlinear, anomalously dispersive material, we performed numerical modeling using the three-dimensional time-dependent generalized nonlinear Schrödinger equation (GNSE) [39, 40] for the amplitude of the field, including all the key physical phenomena, such as dispersion, beam diffraction, Kerr nonlinearity, pulse self-steepening, spatial self-action phenomena, ionization-induced optical

nonlinearities, as well as plasma loss and dispersion. The field evolution equation is solved jointly with the rate equation for the electron density, which includes photoionization and impact ionization.

Numerical modeling is performed using the three-dimensional time-dependent generalized nonlinear Schrödinger equation (GNSE) [39, 40] solved for typical parameters of GaAs—a band gap of 1.4 eV, the Kerr-effect nonlinear refractive index $n_2 \equiv n_{\text{GaAs}} \approx 3 \cdot 10^{-14} \text{ cm}^2/\text{W}$, and the higher order Kerr effect (HOKE) coefficient $n_4 \approx 2 \cdot 10^{-26} \text{ cm}^4/\text{W}^2$ [75]. Dispersion of GaAs was included in the model through the Sellmeier equation [76]. Simulations reproduce all the key features and tendencies in supercontinuum spectra (Fig. 10.10b, e) and pulse shapes (Fig. 10.10c, f), demonstrating the predictive power of our model. In the regime of normal dispersion, the spectral broadening of an ultrashort pulse due to self-phase modulation is limited by dispersion-induced pulse stretching, increasing the pulse width up to 210 fs at $z = 5 \text{ mm}$. In the $N \gg 1$ regime of anomalous dispersion, pulse stretching is suppressed and the field bandwidth grows exponentially at the initial stage of pulse evolution. At later stages, dispersion catches up, compressing the pulse to a pulse width $\tau_s \propto 0.24\tau_0 / (N - 1)$ [67]. For the chosen parameters, the on-axis field intensity at $z = 5 \text{ cm}$ in the regime of anomalous dispersion (0.22 TW/cm^2) is almost an order of magnitude higher than the on-axis field intensity at the same point in the normal dispersion regime.

As one of the key processes in the considered scenario of pulse compression, self-focusing suppresses diffraction-induced beam divergence. The characteristic focusing length of the Kerr-effect-induced nonlinear lens, $l_{\text{sf}} \approx 5.3 \text{ mm}$, is very close to the thickness of the GaAs plate. Such a nonlinear lens, whose focal length is matched with the nonlinear interaction length, keeps the field intensity high, thus helping to maintain the soliton dynamics along the entire propagation path within the GaAs plate despite diffraction. As a result, the spectrum of the mid-IR supercontinuum at the exit interface of the GaAs plate, $z = 5 \text{ mm}$, has a spectrum stretching from 3 to 18 μm (Fig. 10.10e), with its pulse width, $\tau_c \approx 45 \text{ fs}$, corresponding to only 1.2 cycles of the field at $\lambda_0 \approx 7.9 \mu\text{m}$. With an additional chirp compensation in a thin plate of anomalously dispersive material (a 350- μm -thick BaF₂ plate), this self-compressed supercontinuum output can be further compressed, as our simulations show, to a subcycle pulse width of 23 fs.

10.7 Time-Domain Spectroscopy in the Mid-Infrared

When coupled to characteristic, fingerprint vibrational and rotational motions of molecules, an electromagnetic field with an appropriate frequency and waveform offers a highly sensitive, highly informative probe, enabling chemically specific studies on a broad class of systems in physics, chemistry, biology, geosciences, and medicine. The frequencies of these signature molecular modes, however, lie in a region where accurate spectroscopic measurements are extremely difficult because of the lack of efficient detectors and spectrometers. Recent studies [77] show that,

with a combination of advanced ultrafast technologies and nonlinear-optical waveform characterization, time-domain techniques can be advantageously extended to the metrology of fundamental molecular motions in the mid-infrared. In this approach, the spectral modulation of ultrashort mid-infrared pulses, induced by rovibrational motions of molecules, gives rise to interfering coherent dark waveforms in the time domain. These high-visibility interference patterns can be read out by cross-correlation frequency-resolved gating of the field in the visible generated through ultrabroadband four-wave mixing in a gas phase.

Advanced ultrafast laser technologies offer a variety of methods and tools to enhance resonant coupling between electromagnetic radiation and signature molecular modes, opening new avenues for ultrafast spectroscopy [78], chemically specific microscopy [79], quantum chemistry [80], biomedical imaging [81], and standoff detection [82]. In the visible and near-infrared range, resonant frequency combinations [83] and optimized pulse shapes [84] can be used to enhance the coupling between the field and the molecular degrees of freedom through Raman-type processes, enabling time-resolved studies and quantum control of ultrafast processes in molecules [84] and allowing chemically selective bioimaging [81]. In the terahertz range, where electromagnetic radiation can directly probe some of low-frequency characteristic molecular motions, time-domain spectroscopy of molecular bands proves to be advantageous for analytical applications and material characterization [85].

The mid-infrared spectral range is unique, because electromagnetic radiation in this region can resonate with the most intense signature molecular bands. This drastically enhances the coupling between the field and molecular motions, thus offering numerous advantages for molecular spectroscopy [86] and opening routes toward highly sensitive standoff detection [82]. Accurate measurements directly on a mid-infrared signal are, however, difficult because of the lack of efficient detectors and spectrometers for this spectral range. Experiments presented in [77] show that a combination of ultrashort pulses in the mid-infrared and nonlinear-optical methods of pulse characterization are ideally suited for the analysis of fundamental molecular motions, helping confront the long-standing challenges of chemically specific spectroscopy and recognition of molecular modes. We demonstrate that the spectral modulation of an ultrashort mid-infrared pulse induced by molecular rovibrational modes can give rise to high-visibility interference patterns and well-resolved echo revivals in the time domain, which can be read out by means of cross-correlation frequency-resolved optical gating based on broadband four-wave mixing in a gas phase, suggesting a powerful tool for the detection, recognition, and remote sensing of molecular vibrations and rotations.

Ultrashort waveforms with broadband coherent spectra in the mid-infrared play the central role in the time-domain technique for the detection and identification of molecular modes implemented in this work. When the central wavelength of such a waveform is tuned on resonance with a typical molecular rovibrational band, the mid-IR driver, due to its extremely large bandwidth, can interact with the entire manifold of rovibrational transitions, exciting a broadband rovibrational wave packet. As a part of this process, the energy is transferred from the mid-IR driver

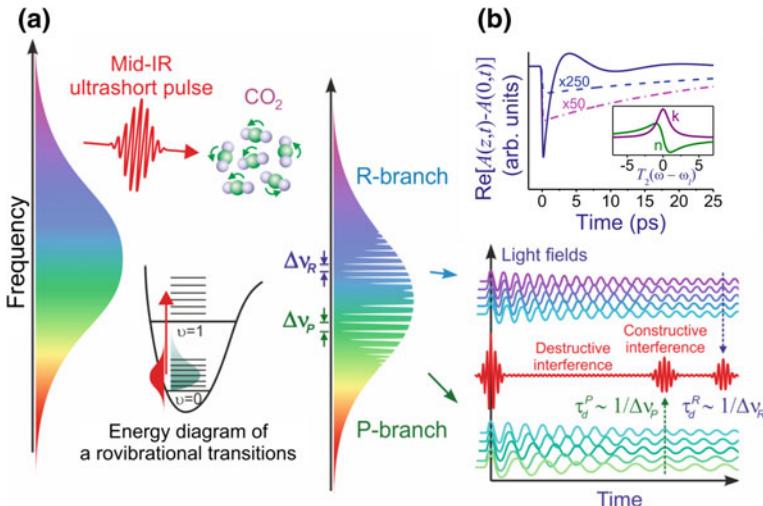


Fig. 10.11 **a** Time-domain spectral interferometry in the mid-infrared. When its central wavelength is tuned on resonance with a typical molecular rovibrational band, the mid-IR driver, due to its extremely large bandwidth, interacts with the entire manifold of rovibrational transitions, exciting a broadband rovibrational wave packet. As a part of this process, the energy is transferred from the mid-IR driver to molecular motion, giving rise to narrowband absorption features in the spectrum of mid-IR radiation at the frequencies of individual molecular modes, separated by spectral intervals $\Delta\nu_P$ and $\Delta\nu_R$ for the P and R rovibrational branches, respectively. In the time domain, these narrowband spectral dips translate into stretched dark field waveforms, as shown in the right panel. Due to the coherence preserved across the entire spectrum of the broadband mid-IR driver, these stretched pulses interfere with each other, giving rise to high-visibility fringes in the waveform of the transmitted mid-IR field (the right panel), with prominent echo field recurrences, observed at delay times $1/\Delta\nu_P$ and $1/\Delta\nu_R$ for the P and R rovibrational branches, respectively. Because the shape of these fringes is fully controlled by the spectrum of molecular rovibrational modes, all the information on a molecule encoded in absorption spectra can be retrieved from mid-IR pulse shapes, thus allowing molecular fingerprints to be read out through a careful analysis of mid-IR waveforms. **b** Dynamics of a dark wave: the envelope of a dark wave, $\text{Re}[A(z, t) - A(0, t)]$, induced by interaction with a molecular transition with a Lorentzian lineshape with $T_2 = 30$ ps and $\Delta\omega_l = 0$ for $\alpha_l z = 0.1$ (pink line), 1 (green line), and 100 (navy line). The input laser profile has a Gaussian envelope and a pulse width of 160 fs. The spectral profiles of the real and imaginary parts n and κ of the complex refractive index $\bar{n} = n - ik$ are shown in the inset. Reproduced from [77]

field to molecular motion, giving rise to narrowband absorption features in the spectrum of the mid-IR driver at the frequencies of individual molecular modes (Fig. 10.11a). In the time domain, these narrowband spectral dips translate into dark field waveforms (Fig. 10.11a), whose phase is shifted by π relative to the phase of the adjacent spectral components and whose duration is much longer than the duration of the incident ultrashort mid-IR driver. Figure 10.1b illustrates the dynamics of one of such dark waveforms, produced by an idealized Lorentzian absorption line. For short propagation paths, $\alpha_l z \ll 1$, with α_l being the absorption

coefficient at the center of the line, the waveform has an exponentially decaying envelope (the dash-dotted line in Fig. 10.11b), as dictated by the Fourier transform of a Lorentzian spectrum. For longer propagation paths, however, propagation effects tend to distort this time-domain map of a molecular mode (see Fig. 10.11b), generally making it more difficult to retrieve the parameters of molecular motions from this dark waveform.

Due to the coherence preserved across the entire spectrum of the broadband mid-IR driver, the dark waves produced by individual absorption lines in a molecular rovibrational band interfere with each other, giving rise to high-visibility fringes in the waveform of the transmitted mid-IR field (the right panel in Fig. 10.11a). Because the shape of these fringes is controlled by the spectrum of molecular rovibrational modes, the information on molecular modes encoded in the absorption spectra can be retrieved from mid-IR pulse shapes, thus allowing molecular fingerprints to be read out through a careful analysis of mid-IR waveforms with allowance for propagation effects (see Fig. 10.11b). Such an analysis is, however, a challenging problem, because of the lack of efficient spectral analyzers and detectors for the mid-IR range. We show below in this paper that this challenge can be confronted using the four-wave mixing (FWM) of mid-IR pulses with well-characterized near-IR reference pulses in a gas phase. This process will be demonstrated to offer a powerful tool for ultrabroadband characterization of complex pulse shapes in the mid-IR, capable of recording ultrafast modulations of mid-IR waveforms, thus allowing unique barcodes of molecular modes to be accurately read out from mid-IR pulses.

In experiments, we use ultrashort mid-IR pulses produced through a difference-frequency mixing [64] of the signal and idler fields delivered by the Ti: sapphire-laser-pumped optical parametric amplification of a broadband seed (Fig. 10.9). The central wavelength of broadband mid-IR fields delivered through this difference-frequency generation process was tunable from 2.7 to 15 μm . Energies above 0.5 μJ were delivered as the DFG output everywhere in the range of wavelengths from 2.7 to 13.6 μm . Pulse characterization of this DFG output was performed using cross-correlation frequency-resolved optical gating (XFROG) based on four-wave mixing (FWM) in a gas medium away from molecular resonances. This technique enables dispersion-free measurements of pulse shapes and spectra of ultrashort mid-IR pulses and allows the spectral and temporal phases of such pulses to be retrieved in a robust and reliable way [64, 77].

In our experiments, this method of time-domain spectroscopy in the mid-infrared was implemented by mixing mid-IR waveforms modulated by molecular motions with reference pulses in a gas medium through four-wave mixing and measuring the frequency-resolved FWM signal as a function of the delay time τ_d between the mid-IR and reference pulses. The 50-fs, 808–810-nm output of the Ti: sapphire laser was used in our experiments as a source of reference pulses, which were mixed with the mid-IR waveforms through the $\omega_{\text{FWM}} = 2\omega_p \pm \omega_d$ FWM process, where the plus and minus signs correspond to sum- and difference-frequency FWM processes, respectively. In contrast to standard XFROG technology [87], which is usually implemented using nonlinear processes in solid-state materials, XFROG

based on FWM in a gas medium is intrinsically broadband, due to a weak dispersion of the gas medium, which helps avoid the stretching of pulses during XFROG measurements and which translates into broadband FWM phase matching, enabling characterization of ultrashort pulses in an ultrabroad spectral range. This version of XFROG is especially advantageous for ultrashort pulses in the mid-infrared as it naturally resolves the material loss problems, which frequently arise in this spectral range for solids. Pulse widths as short as 3.9 to 8 field cycles were detected using this FWM XFROG technique within the range of wavelengths from 3.2 to 10.7 μm , with the smallest pulse-width-to-field-cycle ratio of 3.9 achieved at 5.1 μm , where the DFG output pulse width was found to be 67 fs.

Typical FWM XFROG traces of mid-IR waveforms modulated by the rovibrational modes of carbon dioxide molecules and atmospheric water vapor are shown in Fig. 10.12a–f. The difference between these traces and the XFROG traces of mid-IR pulses tuned off all the molecular resonances is striking. Unlike off-resonance FWM XFROG traces, determined by an appropriate convolution of the mid-IR and reference pulses, the XFROG traces of mid-IR pulses tuned on a resonance with molecular modes exhibit well-resolved fringes along with a complex ring-down structure.

To understand the information encoded in the FWM XFROG traces of mid-IR pulses resonantly driving molecular motions, we need to examine the properties of quantum states of molecules manifested in their rovibrational motions. An antisymmetric stretch of linear molecules, such as carbon dioxide, for example, is coupled to rotations of the molecule, giving rise to characteristic absorption spectra (Fig. 10.12g), which consist of two distinctly different branches of rovibrational transitions with $\Delta J = +1$ (R branch) and $\Delta J = -1$ (P branch) selection rules for the rotational quantum number [88]. Since rotations in rovibrational modes of molecules cannot be decoupled from vibrations, the separation between the individual rovibrational lines $\Delta\nu_J$ slightly varies with J across each of the branches (Fig. 10.12i). However, the mean separation values for these branches can still be meaningfully defined as $\langle\Delta\nu\rangle = c(\sum_J \Delta\nu_J S_J)/(\sum_J S_J)$, where S_J is the absorption line strength, giving, with the use of the reference data for a CO₂ molecule [89], $\langle\Delta\nu\rangle_{\text{P}} \approx 1.82 \text{ cm}^{-1}$ for the P branch and $\langle\Delta\nu\rangle_{\text{R}} \approx 1.29 \text{ cm}^{-1}$ for the R branch.

The coherent waveforms corresponding to the missing, dark spectral components in the spectrum of the driver field, absorbed by P- and R-branch molecular transitions, interfere constructively at delay times $\tau_{\text{P}_m, \text{R}_m} = m/c\langle\Delta\nu\rangle_{\text{P}, \text{R}}$, with m being an integer. This interference gives rise to prominent bright features in the XFROG traces of the mid-IR driver pulse, recurring at $\tau_{\text{P}_1} \approx 18.3 \text{ ps}$ and $\tau_{\text{R}_1} \approx 25.8 \text{ ps}$ (Fig. 10.12c, d, k), corresponding to molecular rovibrational transitions of the P and R branches, respectively. Similar to echo revival signals observed in the emission of rotating molecules, these high-intensity features clearly resolved in our XFROG traces originate from coherent ensembles of molecular rotators. However, unlike rotational echo revivals that show up in the radiative response of a molecular system, the coherent features in the FWM XFROG traces are synthesized through a constructive interference of dark mid-IR waveforms, sampled from the mid-IR driver by molecular motions (Fig. 10.11a). Simulations, performed for an ultrashort

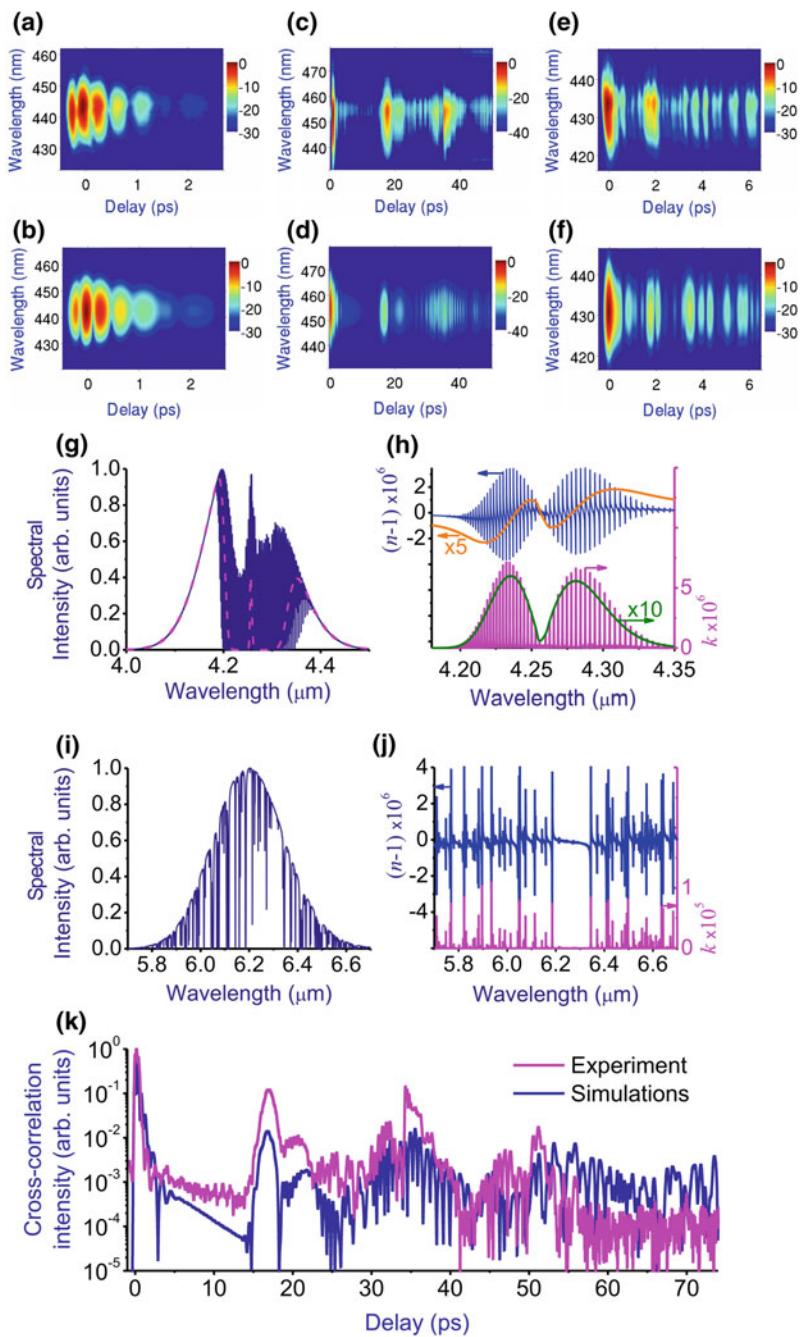


Fig. 10.12 Cross-correlation frequency-resolved optical gating traces (**a–f**) of resonant mid-IR pulses with a central wavelength λ_0 equal to (**a–d**) 4.25 μm and (**e, f**) 6.20 μm and an input pulse width $\tau_0 = 160$ fs transmitted through 1.5 m of atmospheric air: **a, c, e** experiments and **b, d, f** simulations using (1)–(6). The spectrum calculated using only the properly normalized envelopes of the P and R branches of the rotational spectrum of the antisymmetric stretch of a CO₂ molecule is shown by the pink dashed line. **g, h** The spectra of resonant 160-fs mid-IR pulses with a central wavelength of **g** 4.25 μm and **h** 6.20 μm transmitted through 1.5 m of atmospheric air. **i** The spectral profiles of the absorption coefficient and refractive index of the antisymmetric stretch of a CO₂ molecule calculated with the full model of the antisymmetric stretch band of CO₂ (in pink and blue, respectively) and using only a properly normalized envelope of this band, without individual rovibrational lines (in green and orange). **j** The absorption coefficient (pink) and refractive index (blue) of water vapor calculated with the full model of the H–O–H bending mode of a water molecule. **k** The cross-correlation trace of an ultrashort mid-IR pulse with $\tau_0 = 160$ fs and $\lambda_0 = 4.25$ μm transmitted through 1.5 m of atmospheric air: (pink line) experimental data, (blue line) simulations performed with the full model of the rotational spectrum of the antisymmetric stretch of a CO₂ molecule. Reproduced from [77]

mid-IR pulse propagating in a medium with absorption and refraction described with a standard model [88] of rovibrational bands of the antisymmetric stretch in linear molecules (see the Methods section), agree very well with the experimental results (Fig. 10.12c, d, k), confirming the assignment of all the key features in the XFROG and cross-correlation traces measured in our experiments.

Importantly, the coherent peaks induced in the XFROG and cross-correlation traces of a mid-IR pulse by rovibrational molecular transitions are controlled by molecular rotational constants, thus providing a molecule-specific information. This is in contrast with the most intense feature always observed in XFROG and cross-correlation traces at zero delay time between the mid-IR and reference pulses (Fig. 10.12c, d, k), which includes unresolved contributions due to all resonant and nonresonant coherent effects, suggesting no way of retrieving a molecule-specific information. Propagation effects, including absorption in the first place, tend to slightly shift the signature features in XFROG traces and lower their contrast. For our experimental conditions, as the propagation path is increased from 15 to 150 cm the first echo signal related to P-branch rovibrational transitions of CO₂ is shifted from 18.3 to 16.7 ps, while its contrast decreases by two orders of magnitude. The first echo related to the R branch under the same conditions is shifted from 25.8 to 22.3 ps. Still, even with a 150-cm propagation path, the ratio of the maximum intensity of the cross-correlation signal, corresponding to the peak of the P-branch echo signal, to the background intensity exceeds 100 (Fig. 10.12c, d, k), making this echo signal the most intense and the most clearly resolved chemically specific feature in the optical response of the molecule.

Since a water molecule is an asymmetric top, its rotations do not give rise to clearly structured absorption bands. In our scheme of time-resolved measurements, this translates into a much less clearly pronounced structure of both spectra (Fig. 10.12h) and FWM XFROG traces (Fig. 10.12e) of ultrashort mid-IR pulses transmitted through the atmosphere within the 6.27- μm absorption band of water vapor, corresponding to the H–O–H bending mode of a water molecule. Still, even

though the analysis of the experimental data in this case is much more complicated, the experimental XFROG traces can be accurately reproduced (cf. Fig. 10.12e, f) with the use of the standard spectroscopic data (Fig. 10.12j). A complex manifold of intense individual rotational transitions observed in the spectrum of the absorption coefficient and refractive index of atmospheric water vapor (Fig. 10.12j) is seen to give rise to well-resolved bright features in the XFROG traces in Fig. 10.12e, encoding the information on molecular modes, which can be retrieved in certain regimes with an appropriate fitting of these traces (Fig. 10.12e, f).

In full agreement with basic Fourier-transform arguments, individual rovibrational molecular transitions constituting the P and R branches of CO₂ absorption start to show up in XFROG and cross-correlation traces only when the delay time τ_d becomes comparable to τ_{P1} and τ_{R1} (Fig. 10.12c, d, k). This argument helps identify important molecular signatures in cross-correlation and FWM XFROG traces. Specifically, individual rovibrational transitions of a CO₂ molecule translate into a drastic, two-orders-of-magnitude increase in the FWM signal observed within the range of τ_d from 15 to 17 ps, as well as into the second, slightly less pronounced rise of the FWM signal within the range of τ_d from 18 to 21 ps. For shorter delay times, on the other hand, the properties of FWM XFROG traces are largely controlled by the overall shape of P and R rovibrational branches rather than the spectrum of individual rovibrational molecular transitions. This argument is verified by simulations where the full rovibrational spectrum of the antisymmetric stretch of CO₂ (pink line in Fig. 10.12h) is replaced by properly normalized envelopes of the P and R branches (green line in Fig. 10.12h) without their individual spectral components corresponding to transitions with specific values of J . XFROG and cross-correlation traces calculated with the use of this approximation show almost no deviations from the results of calculations performed with the full model of rovibrational spectra of the antisymmetric stretch (Fig. 10.12i). This simplified model is seen to accurately reproduce and explain the most important features in the XFROG traces (Fig. 10.12a–f) and spectra (Fig. 10.12g, h) observed in our experiments, offering a convenient tool to understand the key physical factors and tendencies in the evolution of ultrashort mid-IR pulses within molecular absorption bands.

Thus, the XFROG and cross-correlation traces of mid-IR pulses resonantly coupled to molecular rovibrational modes display signature, molecule-specific features, which show up on distinctly different time scales. These two classes of signatures characterize different aspects of quantized molecular motion. On the one hand, the well-resolved echo signals, appearing in the XFROG and cross-correlation traces of mid-IR pulses on a 10–20-ps time scale provide quantitative information on molecules as quantum rotators. The fringes observed on the finer, subpicosecond time scale, on the other hand, are instrumental in understanding the content of ensembles of quantum molecular rotators and oscillators, allowing the temperature-dependent population distribution over molecular quantum states to be assessed.

To summarize, we have shown that a combination of ultrashort pulses in the mid-infrared and nonlinear-optical methods of pulse characterization are ideally

suited for the analysis of fundamental molecular motions, helping confront the long-standing challenges of chemically specific spectroscopy and recognition of molecular modes. We demonstrate that the spectral modulation of an ultrashort mid-infrared pulse induced by molecular rovibrational modes can give rise to high-visibility interference patterns and well-resolved echo revivals in the time domain, which can be read out by means of cross-correlation frequency-resolved optical gating based on broadband four-wave mixing in a gas phase, suggesting a powerful tool for the detection, recognition, and remote sensing of molecular vibrations and rotations. We have shown that our approach not only offers a tool to retrieve the absorption spectrum, which can be measured with a variety of techniques, most notably, FTIR spectroscopy, but allows the spectrum of the full complex refractive index to be reconstructed.

10.8 Conclusion

Ultrafast optical science is rapidly expanding toward longer wavelengths, into the mid-infrared range, opening the ways toward unique regimes of interaction of high-power coherent electromagnetic radiation with matter, highly sensitive standoff detection, unusual filamentation scenarios and generation of unprecedentedly short field waveforms on the atto- and zeptosecond time scale. The latest breakthroughs in the development of mid-IR sources enabling the generation of sub-100-fs pulses with wavelengths well beyond 3 μm open unique possibilities for systematic studies of optical nonlinearities in the mid-IR within a broad range of field intensities. The first experiments in this direction reveal new nonlinear-optical effects and unusual regimes of nonlinear-optical interactions, suggesting that the extension of standard models of ultrafast nonlinear-optical dynamics to the mid-infrared may be nontrivial. With a rapidly growing toolbox and advanced instrumentation to confront long-standing challenges, the field is heading for new breakthroughs and discoveries in the nearest future. Some of the illustrative material in this chapter appears as a courtesy of the Optical Society of America, American Institute of Physics, and Nature Publishing Group.

References

1. P. Agostini, L.F. DiMauro, Atoms in high intensity mid-infrared pulses. *Contemp. Phys.* **49**, 179–197 (2008)
2. C.I. Blaga, F. Catoire, P. Colosimo, G.G. Paulus, H.G. Muller, P. Agostini, L.F. DiMauro, Strong-field photoionization revisited. *Nat. Phys.* **5**, 335–338 (2008)
3. J.M. Chalmers, P.R. Griffiths (eds.), *Handbook of Vibrational Spectroscopy* (Wiley, 2002)
4. A. Schliesser, N. Picqué, T.W. Hänsch, Mid-infrared frequency combs. *Nat. Photonics* **6**, 440–449 (2012)

5. A. Krier (ed.), *Mid-Infrared Semiconductor Optoelectronics*, vol. 118 (Springer Series in Optical Sciences, 2006)
6. D. Kartashov, S. Ališauskas, G. Andriukaitis, A. Pugžlys, M. Shneider, A. Zheltikov, S.L. Chin, A. Baltuška, Free-space nitrogen gas laser driven by a femtosecond filament. *Phys. Rev. A* **86**, 033831 (2012)
7. G. Andriukaitis, T. Balčiūnas, S. Ališauskas, A. Pugžlys, A. Baltuška, T. Popmintchev, M.-C. Chen, M.M. Murnane, H.C. Kapteyn, 90 GW peak power few-cycle mid-infrared pulses from an optical parametric amplifier. *Opt. Lett.* **36**, 2755–2757 (2011)
8. T. Popmintchev, M.-C. Chen, D. Popmintchev, P. Arpin, S. Brown, S. Ališauskas, G. Andriukaitis, T. Balciunas, O.D. Mücke, A. Pugžlys, A. Baltuška, B. Shim, S.E. Schrauth, A. Gaeta, C. Hernandez-Garcia, L. Plaja, A. Becker, A. Jaron-Becker, M.M. Murnane, H.C. Kapteyn, Bright coherent ultrahigh harmonics in the keV X-ray regime from mid-infrared femtosecond lasers. *Science* **336**, 1287–1291 (2012)
9. J. Weissaupt, V. Juvé, M. Holtz, S. Ku, M. Woerner, T. Elsaesser, S. Ališauskas, A. Pugžlys, A. Baltuška, High-brightness table-top hard X-ray source driven by sub-100-femtosecond mid-infrared pulses. *Nat. Photonics* **8**, 927–930 (2014)
10. A.V. Mitrofanov, A.A. Voronin, D.A. Sidorov-Biryukov, A. Pugžlys, E.A. Stepanov, G. Andriukaitis, T. Flöry, S. Ališauskas, A.B. Fedotov, A. Baltuška, A.M. Zheltikov, Mid-infrared laser filaments in the atmosphere. *Sci. Rep.* **5**, 8368 (2015)
11. P. Panagiotopoulos, P. Whalen, M. Kolesik, J.V. Moloney, Super high-power mid-infrared femtosecond light bullet. *Nat. Photonics* **9**, 543–548 (2015)
12. D. Kartashov, S. Ališauskas, A. Pugžlys, A. Voronin, A. Zheltikov, M. Petrarca, P. Béjot, J. Kasparian, J.-P. Wolf, A. Baltuška, White light generation over three octaves by femtosecond filament at 3.9 μm in argon. *Opt. Lett.* **37**, 3456–3458 (2012)
13. A.V. Mitrofanov, A.A. Voronin, S.I. Mitryukovskiy, D.A. Sidorov-Biryukov, A. Pugžlys, G. Andriukaitis, T. Flöry, E.A. Stepanov, A.B. Fedotov, A. Baltuška, A.M. Zheltikov, Mid-infrared-to-mid-ultraviolet supercontinuum enhanced by third-to-fifteenth odd harmonics. *Opt. Lett.* **40**, 2068–2071 (2015)
14. E.E. Serebryannikov, A.M. Zheltikov, Quantum and semiclassical physics behind ultrafast optical nonlinearity in the midinfrared: the role of ionization dynamics within the field half cycle. *Phys. Rev. Lett.* **113**, 043901 (2014)
15. D. Kartashov, S. Ališauskas, A. Pugžlys, A.A. Voronin, A.M. Zheltikov, A. Baltuška, Third- and fifth-harmonic generation by mid-infrared ultrashort pulses: beyond the fifth-order nonlinearity. *Opt. Lett.* **37**, 2268–2270 (2012)
16. A.A. Lanin, A.A. Voronin, E.A. Stepanov, A.B. Fedotov, A.M. Zheltikov, Multi octave, 3–18 μm sub-two-cycle supercontinua from self-compressing, self-focusing soliton transients in a solid. *Opt. Lett.* **40**, 974–977 (2015)
17. E.A. Stepanov, A.A. Lanin, A.A. Voronin, A.B. Fedotov, A.M. Zheltikov, A solid-state source of subcycle pulses in the mid-infrared, in preparation
18. A.V. Mitrofanov, A.A. Voronin, D.A. Sidorov-Biryukov, A. Pugžlys, E.A. Stepanov, G. Andriukaitis, T. Flöry, S. Ališauskas, A.B. Fedotov, A. Baltuška, A.M. Zheltikov, Mid-infrared laser filaments in the atmosphere. *Sci. Rep.* **5**, 8368 (2015)
19. A.V. Mitrofanov, D.A. Sidorov-Biryukov, A.A. Voronin, A. Pugžlys, G. Andriukaitis, E.A. Stepanov, S. Ališauskas, T. Flöri, A.B. Fedotov, V.Ya. Panchenko, A. Baltuška, A.M. Zheltikov, Subterawatt femtosecond pulses in the mid-infrared range: new spatiotemporal dynamics of high-power electromagnetic fields. *Phys. Usp.* **58**, 89–94 (2015)
20. D. Kartashov, S. Ališauskas, A. Pugžlys, A. Voronin, A. Zheltikov, M. Petrarca, P. Béjot, J. Kasparian, J.-P. Wolf, A. Baltuška, Mid-infrared laser filamentation in molecular gases. *Opt. Lett.* **38**, 3194–3197 (2013)
21. P.B. Corkum, Plasma perspective on strong field multiphoton ionization. *Phys. Rev. Lett.* **71**, 1994–1991 (1993)
22. P. Colosimo, G. Doumy, C.I. Blaga, J. Wheeler, C. Hauri, F. Catoire, J. Tate, R. Chirla, A.M. March, G.G. Paulus, H.G. Müller, P. Agostini, L.F. DiMauro, Scaling strong-field interactions towards the classical limit. *Nat. Phys.* **4**, 386–389 (2008)

23. S. Haessler, J. Caillat, W. Boutu, C. Giovanetti-Texeira, T. Ruchon, T. Auguste, Z. Diveki, P. Breger, A. Maquet, B. Carré, R. Taïeb, P. Salières, Attosecond imaging of molecular electronic wavepackets. *Nat. Phys.* **6**, 200–206 (2010)
24. J. Itatani, J. Levesque, D. Zeidler, H. Niikura, H. Pépin, J. Kieffer, P. Corkum, D. Villeneuve, Tomographic imaging of molecular orbitals. *Nature* **432**, 867–871 (2004)
25. M. Meckel, D. Comtois, D. Zeidler, A. Staudte, D. Pavicic, H. Bandulet, H. Pépin, J. Kieffer, R. Dorner, D. Villeneuve, P. Corkum, Laser-induced electron tunneling and diffraction. *Science* **320**, 1478–1482 (2008)
26. E.E. Serebryannikov, A.M. Zheltikov, in preparation
27. A.M. Zheltikov, Optical phase-space modes, self-focusing, and the wavelength as tunable \hbar . *Phys. Scr.* **90**, 128003 (2015)
28. V.A. Fock, *Electromagnetic Diffraction and Propagation Problems* (Pergamon, London, 1965)
29. J.A. Arnaud, *Beam and Fiber Optics* (Academic Press, New York, 1976)
30. D. Marcuse, *Light Transmission Optics*, 2nd edn. (Van Nostrand Reinhold, New York, 1982)
31. R.W. Boyd, S.G. Lukishova, Y.R. Shen, *Self-focusing: Past and Present. Fundamentals and Prospects* (Springer, New York, 2009)
32. Y.R. Shen, *The Principles of Nonlinear Optics* (Wiley-Interscience, New York, 1984)
33. J.H. Marburger, Self-focusing: Theory. *Prog. Quantum Electron.* **4**, 35–110 (1975)
34. A.M. Zheltikov, Self-focusing and spatial modes in free space and nonlinear waveguides. *Phys. Rev. A* **88**, 063847 (2013)
35. A.B. Fedotov, S.M. Gladkov, N.I. Koroteev, A.M. Zheltikov, Highly efficient frequency tripling of laser radiation in a low-temperature laser-produced gaseous plasma. *J. Opt. Soc. Am. B* **8**, 363–366 (1991)
36. P.B. Corkum, F. Krausz, Attosecond science. *Nat. Phys.* **3**, 381–387 (2007)
37. E.E. Serebryannikov, A.M. Zheltikov, Subcycle waveform generation by nonrecolliding tunneling electron wave packets. *Phys. Rev. A* **90**, 043811 (2014)
38. A. Brown, G. Korn, X. Liu, D. Du, J. Squier, G. Mourou, Self-channeling of high-peak-power femtosecond laser pulses in air. *Opt. Lett.* **20**, 73–75 (1995)
39. A. Couairon, A. Mysyrowicz, Femtosecond filamentation in transparent media. *Phys. Rep.* **441**, 47–189 (2007)
40. L. Berge, S. Skupin, R. Nuter, J. Kasparian, J.-P. Wolf, Ultrashort filaments of light in weakly ionized, optically transparent media. *Rep. Prog. Phys.* **70**, 1633–1713 (2007)
41. J. Kasparian, M. Rodriguez, G. Méjean, J. Yu, E. Salmon, H. Wille, R. Bourayou, S. Frey, Y.-B. André, A. Mysyrowicz, R. Sauerbrey, J.-P. Wolf, L. Wöste, White-light filaments for atmospheric analysis. *Science* **301**(5629), 61–64
42. S. Tzortzakis, B. Lamouroux, A. Chiron, S.D. Moustazis, D. Anglos, M. Franco, B. Prade, A. Mysyrowicz, Femtosecond and picosecond ultraviolet laser filaments in air: experiments and simulations. *Opt. Commun.* **197**, 131–143 (2001)
43. B. Prade, M. Franco, A. Mysyrowicz, A. Couairon, H. Buersing, B. Eberle, M. Krenz, D. Seiffer, O. Vasseur, Spatial mode cleaning by femtosecond filamentation in air. *Opt. Lett.* **31**, 2601–2603 (2006)
44. J.-F. Daigle, A. Jaron-Becker, S. Hosseini, T.-J. Wang, Y. Kamali, G. Roy, A. Becker, S.L. Chin, Intensity clamping measurement of laser filaments in air at 400 and 800 nm. *Phys. Rev. A* **82**, 023405 (2010)
45. D. Mikalauskas, A. Dubietis, R. Danielius, Observation of light filaments induced in air by visible picosecond laser pulses. *Appl. Phys. B* **75**, 899–902 (2002)
46. A.V. Mitrofanov, A.A. Voronin, D.A. Sidorov-Biryukov, G. Andriukaitis, T. Flöry, A. Pugžlys, A.B. Fedotov, J.M. Mikhailova, V.Ya. Panchenko, A. Baltuška, A.M. Zheltikov, Post-filament self-trapping of ultrashort laser pulses. *Opt. Lett.* **39**, 4659–4662 (2014)
47. L. Haizer, I. Bugar, E. Serebryannikov, D. Lorenc, F. Uherek, E. Goulielmakis, A. Zheltikov, Intense Cr:forsterite-laser-based supercontinuum source. *Opt. Lett.* **39**, 5562–5565 (2014)

48. A.A. Voronin, S. Ališauskas, O.D. Mücke, A. Pugžlys, A. Baltuška, A.M. Zheltikov, High-energy-throughput pulse compression by off-axis group-delay compensation in a laser-induced filament. *Phys. Rev. A* **84**, 023832 (2011)
49. S. Drieber, D. Bigourd, N. Fedorov, M. Cornet, M. Arnold, F. Burgi, S. Montant, S. Petit, D. Descamps, E. Cormier, E. Constant, A. Zaïr, Tunable 1.6–2 μm near infrared few-cycle pulse generation by filamentation. *Appl. Phys. Lett.* **102**, 191119 (2013)
50. C.P. Hauri, R.B. Lopez-Martens, C.I. Blaga, K.D. Schultz, J. Cryan, R. Chirla, P. Colosimo, G. Doumy, A.M. March, C. Roedig, E. Sistrunk, J. Tate, J. Wheeler, L.F. DiMauro, E. P. Power, Intense self-compressed, self-phase-stabilized few-cycle pulses at 2 μm from an optical filament. *Opt. Lett.* **32**, 868–870 (2007)
51. R.Y. Chiao, E. Garmire, C.H. Townes, Self-trapping of optical beams. *Phys. Rev. Lett.* **13**, 479–482 (1964)
52. G. Tempea, T. Brabec, Theory of self-focusing in a hollow waveguide. *Opt. Lett.* **23**, 762–764 (1998)
53. V.I. Bespalov, V.I. Talanov, Filamentary structure of light beams in nonlinear liquids. *JETP Lett.* **3**, 307–310 (1966)
54. A.A. Voronin, A.M. Zheltikov, T. Ditmire, B. Rus, G. Korn, Subexawatt few-cycle lightwave generation via multipetawatt pulse compression. *Opt. Commun.* **291**, 299–303 (2013)
55. P.A. Zhokhov, A.M. Zheltikov, Attosecond shock waves. *Phys. Rev. Lett.* **110**, 183903 (2013)
56. A.A. Voronin, A.M. Zheltikov, Subcycle soliton breathers. *Phys. Rev. A* **90**, 043807 (2014)
57. A.B. Fedotov, N.I. Koroteev, M.M.T. Loy, X. Xiao, A.M. Zheltikov, Saturation of third-harmonic generation in a plasma of self-induced optical breakdown due to the self-action of 80-fs light pulses. *Opt. Commun.* **133**, 587–595 (1997)
58. M. Jahjah, W. Jiang, N.P. Sanchez, W. Ren, P. Patimisco, V. Spagnolo, S.C. Herndon, R. J. Griffin, F.K. Tittel, Atmospheric CH₄ and N₂O measurements near Greater Houston area landfills using a QCL-based QEPAS sensor system during DISCOVER-AQ 2013. *Opt. Lett.* **39**, 957–960 (2014)
59. S. Liakat, K.A. Bors, L. Xu, C.M. Woods, J. Doyle, C.F. Gmachl, Noninvasive *in vivo* glucose sensing on human subjects using mid-infrared light. *Biomed. Opt. Express* **5**, 2397–2404 (2014)
60. J. Wegener, R.H. Wilson, H.S. Tapp, Mid-infrared spectroscopy for food analysis: recent new applications and relevant developments in sample presentation methods. *Trends Anal. Chem.* **18**, 85–93 (1999)
61. M.J. Nasse, M.J. Walsh, E.C. Mattson, R. Reininger, A. Kajdacsy-Balla, V. Macias, R. Bhargava, C.J. Hirschmugl, High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams. *Nat. Methods* **8**, 413–416 (2011)
62. M.C. Martin, C. Dabat-Blondeau, M. Unger, J. Sedlmair, D.Y. Parkinson, H.A. Bechtel, B. Illman, J.M. Castro, M. Keiluweit, D. Buschke, B. Ogle, M.J. Nasse, C.J. Hirschmugl, 3D spectral imaging with synchrotron Fourier transform infrared spectro-microtomography. *Nat. Methods* **10**, 861–864 (2013)
63. F. Silva, D.R. Austin, A. Thai, M. Baudisch, M. Hemmer, D. Faccio, A. Couairon, J. Biegert, Multi-octave supercontinuum generation from mid-infrared filamentation in a bulk crystal. *Nat. Commun.* **3**, 807 (2012)
64. A.A. Lanin, A.A. Voronin, E.A. Stepanov, A.B. Fedotov, A.M. Zheltikov, Frequency-tunable sub-two-cycle 60-MW-peak-power free-space waveforms in the mid-infrared. *Opt. Lett.* **39**, 6430–6433 (2014)
65. C.R. Petersen, U. Møller, I. Kubat, B. Zhou, S. Dupont, J. Ramsay, T. Benson, S. Sujecki, N. Abdel-Moneim, Z. Tang, D. Furniss, A. Seddon, O. Bang, Mid-infrared supercontinuum covering the 1.4–13.3 μm molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. *Nat. Photonics* **8**, 830–834 (2014)
66. R. Thomson, C. Leburn, D. Reid (eds.), *Ultrafast Nonlinear Optics* (Springer, New York, 2013)
67. G.P. Agrawal, *Nonlinear Fiber Optics* (Academic, San Diego, 2001)

68. M. Durand, A. Jarnac, A. Houard, Y. Liu, S. Grabielle, N. Forget, A. Durécu, A. Couairon, A. Mysyrowicz, Self-guided propagation of ultrashort laser pulses in the anomalous dispersion region of transparent solids: a new regime of filamentation. *Phys. Rev. Lett.* **110**, 115003 (2013)
69. H. Liang, P. Krogen, R. Grynko, O. Novak, C.-L. Chang, G.J. Stein, D. Weerawarne, B. Shim, F.X. Kärtner, K.-H. Hong, 3-octave supercontinuum generation and sub-2-cycle self-compression of Mid-IR filaments in dielectrics, in *Technical Digest of Meeting on Advanced Solid State Lasers* (Optical Society of America, Washington, 2014), paper ATu5A.4
70. S. Ashihara, Y. Kawahara, Spectral broadening of mid-infrared femtosecond pulses in GaAs. *Opt. Lett.* **34**, 3839–3841 (2009)
71. M. Hemmer, M. Baudisch, A. Thai, A. Couairon, J. Biegert, Self-compression to sub-3-cycle duration of mid-infrared optical pulses in dielectrics. *Opt. Express* **21**, 28095–28102 (2013)
72. A. Pugzlys, P. Malevich, S. Alisauskas, A.A. Voronin, D. Kartashov, A. Baltuska, A. Zheltikov, D. Faccio, in *Technical Digest of Conference on Lasers and Electro-Optics* (Optical Society of America, Washington, 2014), paper FTh1D.3
73. R.A. Kaindl, M. Wurm, K. Reimann, P. Hamm, A.M. Weiner, M. Woerner, Generation, shaping, and characterization of intense femtosecond pulses tunable from 3 to 20 μm . *J. Opt. Soc. Am. B* **17**, 2086–2094 (2000)
74. Y. Nomura, H. Shirai, K. Ishii, N. Tsurumachi, A.A. Voronin, A.M. Zheltikov, T. Fuji, Phase-stable sub-cycle mid-infrared conical emission from filamentation in gases. *Opt. Express* **20**, 24741–24747 (2012)
75. D. Milam, M.J. Weber, A.J. Glass, Nonlinear refractive index of fluoride crystals. *Appl. Phys. Lett.* **31**, 822–825 (1977)
76. M. Bass, C. DeCusatis, J. Enoch, V. Lakshminarayanan, G. Li, C. MacDonald, V. Mahajan, E. Van Stryland, *Handbook of Optics*, vol. 4, 3rd edn. (McGraw-Hill, New York, 2009)
77. A.A. Lanin, A.A. Voronin, A.B. Fedotov, A.M. Zheltikov, Time-domain spectroscopy in the mid-infrared. *Sci. Rep.* **4** (2014)
78. A.H. Zewail, Laser femtochemistry. *Science* **242**, 1645–1653 (1988)
79. B.G. Saar et al., Video-rate molecular imaging *in vivo* with stimulated Raman scattering. *Science* **330**, 1368–1370 (2010)
80. W.W. Warren, H. Rabitz, M. Dahleh, Coherent control of quantum dynamics: the dream is alive. *Science* **259**, 1581–1589 (1993)
81. C.W. Freudiger et al., Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **322**, 1857–1860 (2008)
82. P.N. Malevich et al., Ultrafast-laser-induced backward stimulated Raman scattering for tracing atmospheric gases. *Opt. Express* **20**, 18784–18794 (2012)
83. N. Dudovich, D. Oron, Y. Silberberg, Single-pulse coherently controlled nonlinear Raman spectroscopy and microscopy. *Nature* **418**, 512–514 (2002)
84. J.L. Herek, W. Wohlbaben, R.J. Cogdell, D. Zeidler, M. Motzkus, Quantum control of energy flow in light harvesting. *Nature* **417**, 533–535 (2002)
85. M. Tonouchi, Cutting-edge terahertz technology. *Nat. Photonics* **1**, 97–105 (2007)
86. F. Huth, M. Schnell, J. Wittborn, N. Ocelic, R. Hillenbrand, Infrared-spectroscopic nanoimaging with a thermal source. *Nat. Mater.* **10**, 352–356 (2011)
87. D.T. Reid, P. Loza-Alvarez, C.T.A. Brown, T. Beddard, W. Sibbett, Amplitude and phase measurement of mid-infrared femtosecond pulses by using cross-correlation frequency-resolved optical gating. *Opt. Lett.* **25**, 1478–1480 (2000)
88. G. Herzberg, *Molecular Spectra and Molecular Structure: III. Electronic Spectra and Electronic Structure of Polyatomic Molecules* (Van Nostrand, Princeton, NJ, 1966)
89. L.S. Rothman, R.L. Hawkins, R.B. Watson, R.R. Gamache, Energy levels, intensities, and linewidths of atmospheric carbon dioxide bands. *J. Quant. Spectrosc. Radiat. Transf.* **48**, 537–551 (1992)

Part II

**Historical Works: Single-Photon
and Nonlinear Optical Experiments
in the Pre-Laser Era**

Chapter 11

The First Paper on Experimental Observation of Interference Fringes with Feeble Light (Sir Geoffrey Ingram Taylor)



Svetlana G. Lukishova

Abstract On the editor's selection, this chapter reproduces a brief biography of Sir Geoffrey Ingram Taylor (1886–1975) in Sect. 11.1, and a text of Taylor's paper of the year 1909 “Interference fringes with feeble light” (Sect. 11.2), the first experimental observation of interference fringes with feeble light.

This chapter is devoted to the British scientist Sir Geoffrey Ingram Taylor (1886–1975), a leading figure of the last century in several fields of physics [1–3]. He worked, in addition to optics, in diverse areas such as shock waves, turbulence in the atmosphere and fluids, plastic deformation in solids, physics of explosions and Rayleigh–Taylor instability, electrical activity in thunderstorms (Taylor cones), supersonic aircrafts, etc. During World War II, he was sent to the United States to participate in the Manhattan Project at Los Alamos.

The chapter starts from the article of Sir Brian A. Pippard about Taylor (Sect. 11.1), reprinted with permission of AIP Publishing from *Physics Today* 28, N9, 67 (1975); <https://doi.org/10.1063/1.3069178>. The full text of Taylor's famous 1909 paper “Interference fringes with feeble light”, the first experimental observation of interference fringes with weak light source, is reproduced in Sect. 11.2 from *Proceedings of the Cambridge Philosophical Society* 15, 114–115 (1909) with permission from the Cambridge Philosophical Society.

“Interference fringes with feeble light” was the Taylor's first paper and the only one of his publications devoted to nonclassical optics. At this time Taylor was an undergraduate student and his work was initiated by a Nobel Laureate in Physics (1906), Sir Joseph John Thomson, credited with the discovery and identification of the electron. Thomson earlier introduced in a general form a discontinuity of the light structure (1903), independently from Einstein [4, 5]. Millikan in his Nobel

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road, Wilmot Building, Rochester, NY 14627, USA

e-mail: sluk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

lecture [5] referred to the Thomson–Planck–Einstein concept of localized radiant energy.

Thomson requested Taylor to check whether it would be qualitative modification of a diffraction pattern in the case of light intensity so low that it contains only “few indivisible units” of light energy. Taylor carried out his experiments at different light intensities, but with a constant amount of energy on the photographs by varying the exposure times. Photographs were taken of the shadow of a needle, and the source of light was a narrow slit placed in front of a gas flame. The intensity of the light was reduced by means of smoked glass screens, but with changing the exposure times the total amount of light energy on the photographs was the same in all cases. With highest attenuation the exposure time was 3 months long. The lowest light power in Taylor’s experiment was 5×10^{-6} erg/s ($\sim 10^6$ photons/s) in the region of interference. The results of these experiments showed identical interference-pattern contrast for all incident intensities, even when a single photon struck a particular place of a photographic plate at a time, but with accumulation in time to the same amount of light energy on all photographs.

Statistical properties of the interference patterns in very weak light were investigated by S. I. Vavilov’s group in Russia in 1933, using the human eye as the few-photon detector (see Sect. 14.3.5 of a current book with a translation into English of [6]). True single-photon interference with a real single-photon source was performed for the first time in France in 1986 year by P. Grangier, G. Roger and A. Aspect [7] (see Chap. 1 of this book which also includes some other references after a Taylor paper on interference with feeble light).

- [1] M.P. Brenner and H.A. Stone: Modern classical physics through the work of G.I. Taylor, *Physics Today* **53**, N 5, 30 (2000); <https://doi.org/10.1063/1.883100>.
- [2] G.K. Batchelor: Geoffrey Ingram Taylor, 7 March 1886–27 June 1975. *Biographical Memoirs of Fellows of the Royal Society* **22**, 565–633 (1976), <https://doi.org/10.1098/rsbm.1976.0021>.
- [3] G.K. Batchelor, *The Life and Legacy of G. I. Taylor*, Cambridge University Press (1994).
- [4] R. McCormach: J. J. Thomson and the structure of light, *The British Journal for the History of Science* **3**, No. 4 (Dec.), 362–387 (1967).
- [5] R.A. Millikan: The electron and the light-quant from the experimental point of view, *Nobel Lecture*, May 23 (1924). https://www.nobelprize.org/nobel_prizes/physics/laureates/1923/millikan-lecture.pdf.
- [6] E. Brumberg and S. Vavilov (S. Wawilow): Visuelle Messungen der statistischen Photonenschwankungen (Visual measurements of statistical fluctuations of photons), *Izvestia Acad. Nauk SSSR, Ser. Math. (Bulletin de*

l'Académie des Sciences de l'URSS), OMEN, N 7, 919–941 (1933), in German. Reprinted in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 345–364, in Russian. (See Sect. 14.3 for its translation into English).

- [7] P. Grangier, G. Roger and A. Aspect: Experimental evidence for a photon anticorrelation effect on a beamsplitter, *Europhysics Letters* 1, 173–179 (1986).

11.1 Sir Geoffrey Ingram Taylor: Biography (Excerpts from Sir Brian A. Pippard's Paper in *Physics Today* 28, N9, 67 (1975))

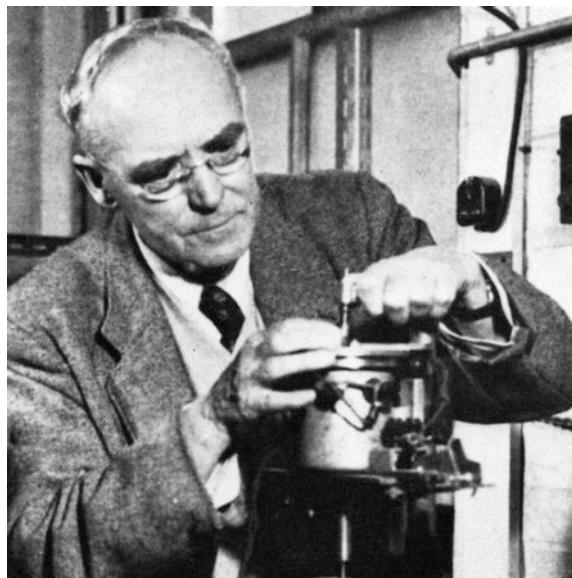
Sir Geoffrey Ingram Taylor (1886–1975) (Fig. 11.1), was one of the great scientists of our time and perhaps the last notable representative of that school of thought that includes Kelvin, Maxwell and Rayleigh, who were physicists, applied mathematicians and engineers—the distinction is irrelevant because their skill knew no such boundaries.

Between 1909 and 1973 he published voluminously, and in a lifetime devoted to research left his mark on every subject he touched and on every one of his colleagues. Although he was no lecturer (some of his attempts to explain difficult matters in simple terms have become legend), his outgoing manner and complete lack of pomposity conveyed, as no formal exposition could have done, the enthusiasm and intuitive understanding that informed all his work.

His first paper, on interference fringes photographed in very weak light, was his only excursion into the world of quanta and was undertaken, according to his own story, because the exposure time of hundreds of hours enabled him to fit in an already-arranged sailing voyage. Afterwards problems of continuum mechanics mainly attracted him, especially fluid dynamics and the plastic deformation of crystals. It is strange to note that his pioneering contribution to the initiation of dislocation theory in 1934 very nearly represented the end of his researches in this field, because immediately after he became wholly immersed in the statistical theory of turbulence. His great gift for combining mathematical analysis with relatively simple experiments, each fertilizing the other, is shown here at its most powerful. It was a joy to see him demonstrate some phenomenon in fluid flow with bits and pieces seemingly gathered from the scrap box and assembled in the kitchen sink. This was not the economy of a lazy man—quite the opposite; such demonstrations were an essential part of the equipment of an astonishingly agile and exact mind, enough to give a hint on which to build an analytical description.

And this ability to see interesting problems in the observations of daily life stayed with him to the end—water-bells, paint-rollers, the swimming of eels and sperm, peeling of adhesive tape, disintegration of charged drops—all these came under his imaginative scrutiny after he had reached the age of 65, and all served to demonstrate how a vigorous mind need not seek challenges among the expensive

Fig. 11.1 G. I. Taylor
(courtesy of the Cavendish
Laboratory, University of
Cambridge, from Emilio
Segrè Visual Archives)



equipment of a modern laboratory. To be sure, none of these later investigations broke new ground, but are the last outpouring of a fertile genius who, as much as any other, kept the flame of classical physics burning brightly when it could well have been extinguished.

Taylor spent the greater part of his life at Trinity College, Cambridge and at the Cavendish Laboratory, where he was a Royal Society Professor. Honorary doctorates from all around the world came his way and provided opportunities for travel and exploration that formed his private relaxation. He was admitted to the Order of Merit in 1969.

To his many friends he was a continual inspiration, at once a profound thinker and, it seemed, a truly happy man.

11.2 Reprint of the First Paper on Experimental Observation of Interference Fringes with Feeble Light [from *Proceedings of the Cambridge Philosophical Society* 15, 114–115 (1909)]

Interference fringes with feeble light [By G. I. Taylor, B. A., Trinity College. (Communicated by Professor Sir J. J. Thomson, F. R. S.). (Read 25 January 1909)].

The phenomena of ionisation by light and by Röntgen rays have led to a theory according to which energy is distributed unevenly over the wave-front (J. J. Thomson, *Proc. Camb. Phil. Soc.* XIV. p. 417, 1907). There are regions of

maximum energy widely separated by large undisturbed areas. When the intensity of light is reduced these regions become more widely separated, but the amount of energy in any one of them does not change; that is, they are indivisible units.

So far all the evidence brought forward in support of the theory has been of an indirect nature; for all ordinary optical phenomena are average effects, and are therefore incapable of differentiating between the usual electromagnetic theory and the modification of it that we are considering. Sir J. J. Thomson however suggested that if the intensity of light in a diffraction pattern were so greatly reduced that only a few of these indivisible units of energy should occur on a Huygens zone at once the ordinary phenomena of diffraction would be modified. Photographs were taken of the shadow of a needle, the source of light being a narrow slit placed in front of a gas flame. The intensity of the light was reduced by means of smoked glass screens.

Before making any exposures it was necessary to find out what proportion of the light was cut off by these screens. A plate was exposed to direct gas light for a certain time. The gas flame was then shaded by the various screens that were to be used, and other plates of the same kind were exposed till they came out as black as the first plate on being completely developed. The times of exposure necessary to produce this result were taken as inversely proportional to the intensities. Experiments made to test the truth of this assumption shewed it to be true if the light was not very feeble.

Five diffraction photographs were then taken, the first with direct light and the others with the various screens inserted between the gas flame and the slit. The time of exposure for the first photograph was obtained by trial, a certain standard of blackness being attained by the plate when fully developed. The remaining times of exposure were taken from the first in the inverse ratio of the corresponding intensities. The longest time was 2000 h or about 3 months. In no case was there any diminution in the sharpness of the pattern although the plates did not all reach the standard blackness of the first photograph.

In order to get some idea of the energy of the light falling on the plates in these experiments a plate of the same kind was exposed at a distance of two metres from a standard candle till complete development brought it up to the standard of blackness. Ten seconds sufficed for this. A simple calculation will shew that the amount of energy falling on the plate during the longest exposure was the same as that due to a standard candle burning at a distance slightly exceeding a mile. Taking the value given by Drude for the energy in the visible part of the spectrum of a standard candle, the amount of energy falling on 1 cm^2 of the plate is 5×10^{-6} ergs/s and the amount of energy per cubic centimetre of this radiation is 1.6×10^{-16} ergs.

According to Sir J. J. Thomson this value sets an upper limit to the amount of energy contained in one of the indivisible units mentioned above.

Chapter 12

First Experiments on Measuring Light Pressure I (Pyotr Nikolaevich Lebedev)



Anatoly V. Masalov

Abstract This chapter is devoted to the first experiment of measuring light pressure by P. N. Lebedev (1866–1912). In Sect. 12.1, A. V. Masalov describes the details of this experiment. Section 12.2 contains the translation into English one of the first Lebedev's papers of 1901. It is a translation from the Journal of the Russian Physico-chemical Society (*Zhurnal Russkogo Fiziko-Khimicheskogo obschestva* (Zh.R.F.Kh.O.)), Ser. Physics 33 (1), 53–75 (1901). Журнал Русского физико-химического общества (Ж.Р.Ф.Х.О.), часть физическая, 33(1), 53–75 (1901). The same paper was published in German: P. Lebedew: *Ann. d. Phys.*, 4, Is. 6, 433–458 (1901).

12.1 P. N. Lebedev—First Experiment on Measuring Light Pressure

In 1901 the leading European physical journal had published the paper by P. Lebedev (Fig. 12.1) “Untersuchungen über die Druckkräfte des Lihtes” (“Researches on the pressure forces of light”) [1]. Russian equivalent of paper [1] was published in the Journal of Russian Physical and Chemical Association [2]. The final conclusion of these papers was the following: “... within the limits of error the existence of light pressure of Maxwell and Bartoli are *quantitatively* confirmed.”

The matter concerns the experimental confirmation of the formula for the radiation pressure $P = \frac{E}{c}(1 + \rho)$, which is expressed in terms of the luminous flux E and the reflectivity of the illuminated surface ρ (c is the speed of light). A detailed description of the measurement method and the results given in papers [1, 2] was preceded by a report presented by P. Lebedev at the First International Congress of Physics in Paris (August 1900) [3].

A. V. Masalov (✉)

Lebedev Physical Institute, Leninsky Prospekt 53, Moscow 119991, Russia
e-mail: masalov@sci.lebedev.ru

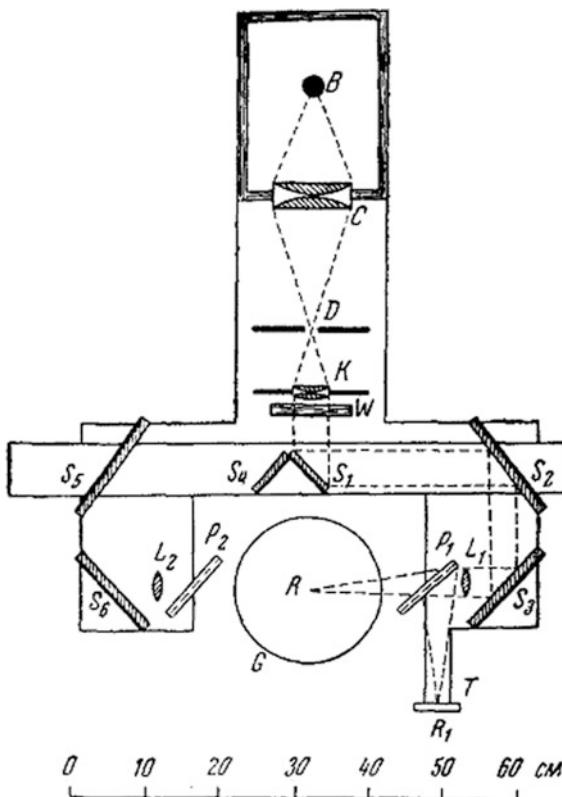
Fig. 12.1 P. N. Lebedev
(courtesy of P. N. Lebedev
Physical Institute, Moscow,
Russia)



Lebedev's measurements (Fig. 12.2) were performed with an arc lamp that gave light intensity few times greater than that of the sunlight on the surface of the Earth. To measure the pressure forces with sensitivity at the level of 10^{-6} dyne he built the setup where the light irradiated one vane of two, suspended on a quartz fiber in a vacuum jar (= torsion balance in a vacuum). Full description in English of setup components, as well as the results of measurements, equivalent to [1, 2], is absent, there is only a brief summary in [4].

When building the setup of such a high sensitivity Lebedev had to eliminate two factors that disturb the light pressure observation. Both factors are related to the impact of residual gas in a vessel with a torsion balance. The first factor is the radiometric force acting on the vane due to the temperature difference between illuminated and backward surface. In the rarefied gas the temperature difference between surfaces causes the difference of gas pressure that can be present for some time, which causes the radiometric force. The action of radiometric force was clearly demonstrated in the experiments of Crookes [5]. With moderate vacuum $> 10^{-3}$ bar and sunlight illumination the radiometric force is several orders of magnitude greater than that of light pressure. The second factor is the vertical convection of gas heated by illuminated vane; if the vane surfaces deviates from the vertical plane the convection produce torque. In Lebedev setup both factors were removed by high vacuum (i), by the use of thin metal vanes (ii) and by alternately illuminating the vane on the one or the other side (iii). Obviously, the decrease in residual gas pressure in jar with a torsion balance reduces both radiometric and convection forces. To achieve ultimate vacuum Lebedev used a standard (for the time) mercury pump which ensures vacuum of 10^{-4} bar; further vacuum

Fig. 12.2 Lebedev's experimental setup for measuring the radiation pressure force (top view):
B—arc lamp, **C**—condenser,
D—diaphragm 4 mm,
K—collimating lens,
W—color filter, S_1 ,
 S_4 —mirrors on a moving plate (at left plate position the light beam is directed right, and vice versa), S_2 , S_3 , S_5 ,
 S_6 —fixed mirrors, L_1 ,
 L_2 —lens, **G**—glass jar with torsion balance, P_1 , P_2 —glass plates, **R**—image of diaphragm **D**, R_1 —image of diaphragm **D**,
T—thermocouple



improvement was achieved due to the diffusion of mercury vapor—the method invented by Lebedev.

Vanес of torsion balance (Figs. 12.3, 12.4 and 12.5) Lebedev had made of different materials. In particular, as an absorbing material he used blackened platinum foil of thickness 0.02 and 0.1 mm. Lebedev expected that the thin foil would provide such a rapid heat transfer between heated and shaded surfaces, that the temperature difference would be negligible together with radiometric force. A thick foil will serve as an indicator of the appearance of radiometric forces.

The method of alternately illuminating the vane on one or other side allowed Lebedev to eliminate the influence of gas convection. During measurements Lebedev registered the mean point of balance oscillation in two cases of illumination. The light pressure force shifts these points oppositely, while the convection force shifts them equally in one direction. Thus, the difference between mean points of oscillation in two cases of illumination is a measure of light pressure. During the observations Lebedev found that at limit vacuum achieved in his setup there is no drift of “zero” position of the torsion balance; he used this indicator as a condition for successful measurements. In addition, measurements with thick and thin platinum foil did not differ from each other, which indicated the absence of disturbance

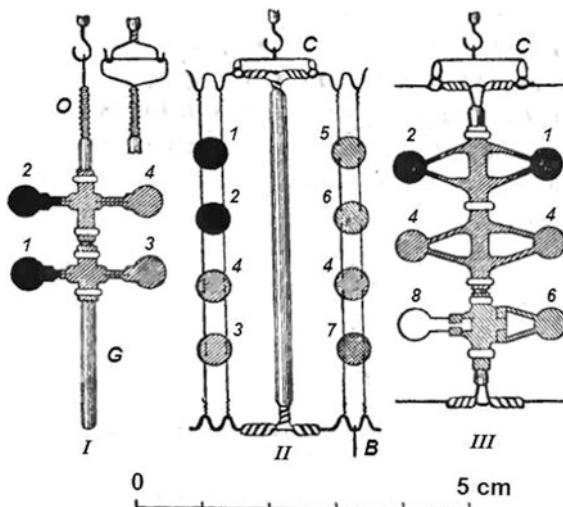


Fig. 12.3 Vanes (5 mm dia.) used in Lebedev's measurements: 1—blackened platinum (about 0.1 mm thick), 2—blackened platinum (about 0.02 mm thick), 3—platinum mirror 0.1 mm thick, 4—platinum mirror 0.02 mm thick, 5—aluminium mirror 0.1 mm thick, 6—aluminium mirror 0.02 mm thick, 7—nickel mirror 0.02 mm thick, 8—mica < 0.01 mm thick

Vane Used	RADIATION PRESSURE	
	Observed	Computed
Black, platinum plated	1.1	1.0
Bright platinum	1.8	1.6
Bright aluminium	2.0	1.8
Bright nickel	1.4	1.6

Fig. 12.4 Results of radiation pressure measurements

by radiometric forces. After the necessary calibrations (torsion constant of fiber, reflectivity of vane surfaces) and measurements, Lebedev obtained the following results [4] (Fig. 12.4):

Errors of the measurements Lebedev estimated at 20%. Lebedev had formulated conclusions of the work done:

1. Light beam produces a pressure on reflecting and on absorbing surfaces, and this pressure is not related to yet known convectional and radiometric forces which are caused by heating of residual gas.

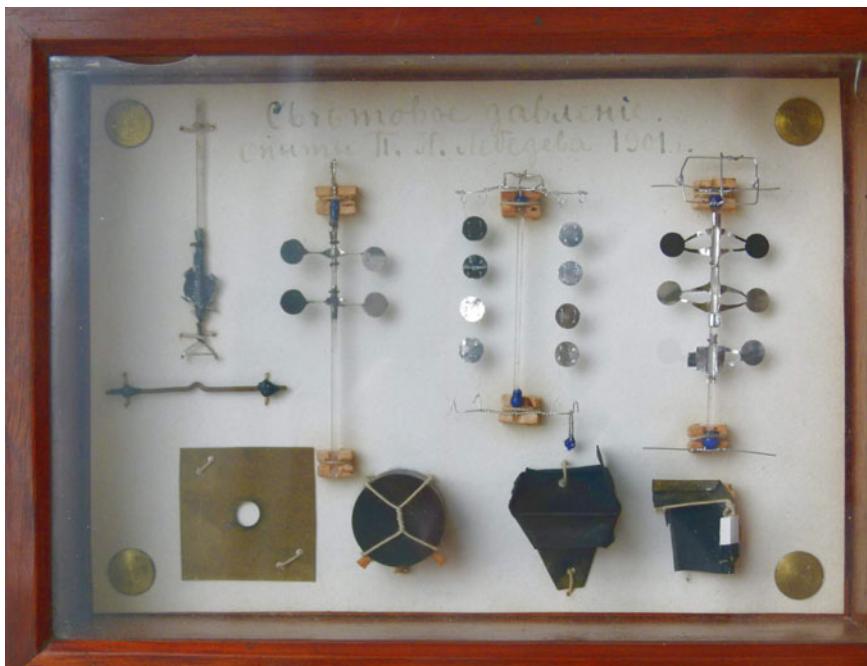


Fig. 12.5 Various vanes and other components made by P. N. Lebedev for his setup (maintained in the Lebedev Physical Institute, Moscow). Upper legend: “Light pressure, experiments of P. N. Lebedev, 1901.” Internal frame size: 12 × 17 cm. (Courtesy of A. V. Kraisky)

2. Light pressure is proportional to light power and does not depend on light color.
3. Experimental values of light pressure are in a quantitative agreement within the limits of error with radiation pressure by Maxwell and Bartoli.

It is expedient and instructive to note how Lebedev has evaluated the resulting errors of his measurements. He pointed out three groups of errors:

1. Measurement error of the pressure force that includes torsion coefficient of fiber, the error of distance from the setup to the scale where the reflection from the mirror on torsion balance indicated the angle of rotation, the error of radius from fiber to center of vane. The latter error was 5% and dominated in this group. Lebedev estimated the overall error at 8%.
2. The error in calculating the absolute value of the light power on vane that includes an uncertainty of the water equivalent of calorimeter, calorimeter temperature measurement error and discrepancy between vane square and calorimeter input aperture. For example, two calorimeters testified the same luminous flux showed 5% difference. Lebedev evaluated the overall error at 7%.
3. The error in determining the reflection coefficients of vanes used. Because of surface imperfections of vanes Lebedev evaluated their measurement error at 10%.

The final error Lebedev estimated at 20% worrying not to give an unjustified underestimation of errors. In an experiment with unexampled sensitivity of measurements and with the demonstration of the new effect such a position of the author deserves respect.

At the same time in 1900/1901 when P. Lebedev was measuring the light pressure, E. F. Nichols and G. F. Hull worked at light pressure measurements in Dartmouth College (USA). They worked independently and elaborated their own setup to measure the light pressure. In 1901, E. F. Nichols and G. F. Hull published a preliminary report on the method they used for these measurements and on their first results [6]. In conclusion they claimed the relation between the measured radiation pressure and the radiation pressure estimated on measured light power as 1.05:1.34 or as 78:100. Later in 1903, E. F. Nichols and G. F. Hull improved precision of measurements and presented detailed description of the experimental setup and measurement results in paper [7], where the agreement between measured pressure and estimations were stated as 1% (regarding the accuracy of their measurements, see Chap. 13).

After the talk of P. Lebedev at the First International Congress of Physics in Paris in August 1900, renowned British physicist Lord Kelvin said to another participant—the famous Russian scientist K. A. Timiryazev [8]: “Maybe you know that all my life I struggled against Maxwell not admitting his light pressure, and now yours Lebedev make me to give up in front of his experiments.”

In the Archives of the Academy of Science of USSR there are several letters of scientists from abroad, who responded to Lebedev’s results on light pressure measurements. Few letters were translated into Russian. One letter from F. Paschen to P. Lebedev (November 18, 1901, Tübingen) is quoted here [9]: “... I am also convinced that the forces of light pressure does exist. And I believe that your work is at least the first important evidence of the existence of these forces. It is utterly improbable that good agreement with theoretical predictions is accidental. I am not surprised that you allow discrepancy up to 20%, and I think that the question is not in the precise agreement. You have shown the way to achieve this important result. Let other followers to work in order to make your method more delicate. There will be no lack of such people.”

For Lebedev’s biography and full publication list see [10] and [11].

References

1. P. Lebedew, “Untersuchungen über die Druckkräfte des Lihtes” (“Investigation of light pressure”): *Annalen der Physik*, **6**, 433–458 (1901).
2. П.Н. Лебедев, “Опытное исследование светового давления” (P.N. Lebedev, “Experimental examination of light pressure”). *Ж.Р.Ф.Х.О. (ч. физ.)*, **33**(1), 53–75 (1901). English translation of this paper is reproduced in this book from web.ihep.su/dbserv/compas/src/lebedev01/eng.pdf.
3. P. Lebedew, “Les forces de Maxwell-Bartoli dues à la pression de la lumière” (P. Lebedev, “The forces of Maxwell-Bartoli of light pressure”). *Rapports présentés au Congrès Internationale de Physique*, **2**, 133–140 (1900). See online “The Scientific Heritage of Russia”: <http://books.e-heritage.ru/book/10077976>. The same in Russian: П.Лебедев, “Максвелло-Бартолиевские силы давления лучистой энергии”. *Ж.Р.Ф.Х.О (ч.физ.)*, **32**(1), 211–217 (1900).

4. P. Lebedew, “An experimental investigation of the pressure of light.” *Astrophys. Journal*, **15**, 60–62 (1902).
5. W. Crookes, “On attraction and repulsion resulting from radiation.” *Philos. Trans. R. Soc. London*, **164**, 501–527 (1874).
6. E.F. Nichols and G.F. Hull, “A preliminary communication on the pressure of heat and light radiation.” *Phys. Rev.*, Series I, **13**, 307–320 (1901).
7. E.F. Nichols and G.F. Hull, “The pressure due to radiation.” *Phys. Rev.*, Series I, **17**, 26–50 (1903). *Ibid.* **17**, 91–104 (1903).
8. С.И. Вавилов, в книге *Люди русской науки*, т.1. “Пётр Николаевич Лебедев” (S.I. Vavilov, in *People of Russian science*, v.1. “Petr Nikolaevich Lebedev”). ОГИЗ, М.-Л., 1948, стр. 244.
9. In *Научное наследство*, т. I, под ред. С.И. Вавилова и др. (*Scientific Heritage*, v.1. Eds. S.I.Vavilov et al.). Изд. АН СССР, М.-Л., 1948, стр. 569–570.
10. E.I. Pogrebysskaya: Piotr Nikolaevich Lebedev, *Journ. of Russian Laser Research* **37**, N 5, 418–424 (2016).
11. A.M. Lukomskaya, *Piotr Nikolaevich Lebedev*, Bibliography, Soviet Academy of Sciences, Commission of History of Physics and Mathematics, Soviet Academy of Sciences Publ., Leningrad (1950), 185 pages, in Russian.

12.2 “Experimental Examination of Light Pressure”. Original Paper of P. N. Lebedev on First Experimental Examination of Light Pressure

*Translation from Russian into English was provided by V. O. Soloviev
Translation editing was provided by A. V. Masalov*

Publication is permitted by the Nauka Publisher, Moscow, Russia.

Explicating the basic standings of the electromagnetic theory of light Maxwell (1873) has paid attention also to those forces which arise to us as ponderomotive forces in every magnetically- or electrically-polarized medium: necessity of existence of these forces inevitably follows from his theory in any bundle of rays also, and Maxwell¹ tells us:

In a medium in which waves are propagated there is a pressure in the direction normal to the wave, and numerically equal to the energy contained in unit of volume.

¹J. C. Maxwell, Treatise on electricity and magnetism, § 792.

The further substantiation of these Maxwell forces of pressure of electromagnetic waves we discover at O. Heaviside,² H. A. Lorentz,³ E. Cohn,⁴ and D. Holdhammer.⁵

A. Bartoli (1876)⁶ has come to an identical conclusion following completely different way and, probably, being not informed of the ray property indicated by Maxwell. Bartoli specifies cyclic processes, which should enable by means of moving mirrors to transfer a radiant energy from a more cold body to a warmer one, and evaluates the work, which should be done in this case according to the second law of thermodynamics. Necessity to expend a work by moving a mirror towards the impinging beam forces to assume, that the impinging beam presses on a mirror. Bartoli has calculated a value of this pressure; the effect obtained by him completely coincides with the effect obtained by Maxwell.

Boltzmann⁷ has followed along the path indicated by Bartoli at evaluations of pressure of beams, and then Prince B. B. Galitzine⁸ and Guillaume,⁹ while Drude¹⁰ has extended this method onto the absolutely black body.

If a parallel bundle of beams impinges steeply on a flat surface, the amount of Maxwell-Bartoli pressure is determined by the amount of energy E , impinging per second, by reflectivity of a surface and by velocity v of the beam propagation; then

$$p = \frac{E}{v}(1 + \rho),$$

where ρ is in the range between 0 for the absolutely black surface and 1 for the absolutely reflecting surface.

The value of this beam pressure is rather small. Both Maxwell and Bartoli have calculated that the Sun rays, impinging steeply on a flat surface of 1 m^2 , should yield pressure, which in a case of a black surface is equal to 0.4 mg, and in case of a mirror –0.8 mg.

The assumptions that the beams of light should yield pressure, were expressed already much earlier. So, Kepler (1619), trying to explain the specific shape of comet tails, for the first time has stated an idea, that this shape is stipulated by pressure of solar beams on particles of substance of tails; this guess was in the complete accordance with a outflow hypothesis prevailed that time and has found

²O. Heaviside, Electromagnetic Theory **1**, 334 (London, 1893).

³H. A. Lorentz, Versuch einer Theorie der elektromagnetischen und optischen Erscheinungen in bewegten Körpern, page 29 (Leiden, 1895).

⁴E. Cohn, Das elektromagnetische Feld, page 543 (Leipzig, 1900).

⁵D. Holdhammer, Ann. d. Phys. **4**, 834 (1901).

⁶A. Bartoli, Exner's Rep. d. Physik **21**, 198 (1884) German translation from Nuovo Cimento **15**, 195 (1883).

⁷L. Boltzmann. Wied. Ann. **22**, pages 33, 291, 616 (1884).

⁸B. Galitzine, Wied. Ann. **47**, 479 (1892).

⁹Ch. Ed. Guillaume, Archives des Sciences phys. et nat. de Genève **31**, 121 (1894).

¹⁰P. Drude, Lehrbuch der Optik, page 447 (Leipzig, 1900).

hot support from Longomontanus (1622)¹¹. The same effect has inspired L. Euler (1746)¹² to assign pressing forces to a light beam, and he has made attempt to justify them theoretically, viewing a light wave (according to Huygens) as longitudinal oscillations.

De Mairan (1754)¹³ has undertaken together with Du Fay the first rather interesting experiments to be convinced of validity of the guesses mentioned above, but he should leave them, as the convectional currents in an ambient air hindered the observation of a guessed effect. If to take into consideration those resorts, which could be arranged by the experimenter in XVIII century, De Mairan experiments deserve the greatest surprise. The similar experiments were undertaken then by Fresnel (1825)¹⁴, who have been stopped by the same difficulties; detailed study of appearances having here a place, has lead W. Crooks¹⁵ to discovery of radiometric forces.

Maxwell-Bartoli forces of beam pressure can in due course receive a great value in problems of physics and astronomy, that is why the experimental examination of these forces is even more advisable, as their theoretical substantiations both according to Maxwell and to Bartoli are based on particular partial properties of absorbing and reflecting surfaces, and consequently there can be a problem, whether the forces of pressure are really stipulated *only* by these partial properties of surfaces in a case of *light rays* also. This problem can be solved only through extra examinations; the most direct way is the immediate experiment.

Attempts by F. Zöllner¹⁶ and Bartoli (cited above, page 205), made in this direction have not given positive results; that is why I also have undertaken the following experimental examination of light pressure.¹⁷

I. Preliminary Experiments

In his textbook Maxwell (§ 793) tells us:

It is probable that a much greater energy of radiation might be obtained by means of the concentrated rays of the electric lamp (than solar light). Such rays falling on a thin metallic disk, delicately suspended in a vacuum, might perhaps produce an observable mechanical effect.

¹¹See below in de Mairan, page 355–356.

¹²L. Euler, Histoire de l'Academie de Berlin **2**, 121 (1746).

¹³De Mairan, Traité physique et historique de l'Aurore Boréale (Seconde Edition), page 371 (Paris, 1754).

¹⁴A. Fresnel, Ann. de Chimie et de Phys. (2) **29**, 57, 107 (1825).

¹⁵W. Crooks, Philos. Transact. of the R. S. of London **164**, 501 (1874); in this article there is a list of references concerned here.

¹⁶F. Zöllner, Pogg. Ann. **160**, 154 (1877).

¹⁷A draft Report about this examination was made by me on the First International Congress of Physics in Paris (in August 1900); the translation of contribution is published in Zhurnal Russkogo Fiziko-Khimicheskogo Obschestva (Fizika) **32** (1), page 211, (1900).

When I started with the experiments, I supposed that the arrangement indicated by Maxwell does not lead to the goal as F. Zöllner¹⁸ has already failed on this way; also he has “paid attention to the circumstance, that numerical quantity (of the light pressure), theoretically predicted by the Maxwell, is approximately 100 000 times less than observed by Crooks in one special case.”¹⁹ If it was even possible to hope to reduce in a very considerable measure these secondary radiometric forces, nevertheless, it seemed to me, that only such experiment could have the decisive meaning, in which it would be possible to cancel somehow the activity of these forces.

At examination of radiometric forces Schuster²⁰ has shown, that they are interior forces of a radiometer; Righi²¹ confirmed this result by a very refined experiment: “I have arranged so,—Righi tells us,—that the radiometer floated on a surface of water upside down; the glass cap of a propeller laid thus on that tube, which is ordinary retains a rotated rod of a propeller in a vertical standing. Due to this there was frictional force, not allowing a gyration of a propeller. When I now have guided on a vane of a propeller a strong beam of a light, I could not detect slightest gyration (of the radiometer).”

Both Bertin and Garbe²² came to the same conclusion in repeating this experiment.

Wishing to detect in experiment Maxwell-Bartoli forces of light pressure, I have taken advantage of Righi’s arrangement in such a way: a mica plate, bent into a cylinder, was fixed between two circles which have been cut out from a thin nickel leaf. The cylinder served as a body of the radiometer; inside it there was a vane immobilely fastened with it. This radiometer was suspended on a glass fiber inside the evacuated glass bulb. When I guided a light of an arc lamp onto the vane, I permanently observed²³ deviations, which were of the same order as ones evaluated according to Maxwell-Bartoli.²⁴

When, during these preliminary experiments, I began to study for comparison the forces acting just on the vane, without a mica shell, I had found, that the

¹⁸F. Zöllner, cited above, page 155.

¹⁹F. Zöllner has put too small energy for a radiation of candle in basis of the calculation. If compare radiometric forces observed by E. Nichols (Wied. Ann. **60**, 405 (1897)), with those forces of pressure evaluated according to Maxwell and Bartoli from Angström (Wied. Ann. **67**, 647 (1899)) data, concerning radiation of a new candle, the relation gained is about 10 000.

²⁰A. Schuster, Phil. Mag. (5) **2**, 313 (1876).

²¹A. Righi, the literal translation is given at Bertin et Garbe, see below.

²²Bertin et Garbe, Ann. de Chim. et de Phys. (5) **11**, 67 (1877).

²³If Righi and also Bertin and Garbe have not noted *any* Maxwell-Bartoli forces, it follows *extremely* that their arrangement calculated for much more radiometric forces, was insufficiently sensitive to measure forces of light pressure.

²⁴Results of these preliminary experiments were reported on May 17, 1899 at session Société Vaudoise in Lausanne (Arch. des Sc. phys. et nat. Genève **8**, 184 (1899)). The casual circumstances have interfered and prevented opportune appearance of a detailed note planned, and it has remained not printed.

radiometric forces, observed at it, were far from being reach the value specified by F. Zöllner. The perturbation induced by them, appears even less than the perturbation stipulated by a convection. The last is exhibited in a very strong degree at the rather large sizes of a glass bulb of a radiometer. Therefore I have left this method and have gone to other experiments, which I provided on a prime method indicated by Maxwell.

II. An Arrangement of Experiments and Devices

Though Maxwell arrangement of experiment is rather simple, it meets, however, two essential difficulties stipulated, first, *by convectional flows*, and second—*by radiometric forces*. These secondary forces considerably diminish at the highest rarefactions, but nevertheless it is necessary to consider them when measuring the light pressure.

The origin of convectional forces is stipulated by the fact that when heating up a vane of the device by impinging beams, the adjacent strata of gas are heated simultaneously and the uprising flux is formed; if the plane of a vane is slightly canted in relation to a vertical, then the uprising flux forces a vane to move. The direction and the value of this displacement depend *only* on a degree of heating up and *do not depend* on a direction, on which the heating beams impinge. These forces can be eliminated at measurements, forcing beams from the same source to impinge alternately with one or the other hand of vane.

As to radiometric forces, they were reduced in my experiments up to the possible minimum due to taking a rather large glass bulb²⁵ ($D = 20$ cm). And all beams which could be absorbed by walls of the bulb were eliminated²⁶ by the relevant light filter, vanes were made of thin metal, so the odds of temperatures of both surfaces were small whenever possible, and rarefaction²⁷ was entered (through the mercury pump and its subsequent cooling by a cooling mixture) up to highest possible rate.

When the radiometric forces are small, the correction at measurement light pressure due to them can be calculated on the following bases: the radiometric forces are caused by an odds of temperatures between irradiated and not irradiated vane surfaces, and for two isometric vanes made of an identical material and surfaces having identical properties, these forces are directly proportional to *thicknesses*²⁸ of vanes. If we shall observe simultaneously *two* identical vanes having very considerable odds of thickness, we can *calculate*, how great would be

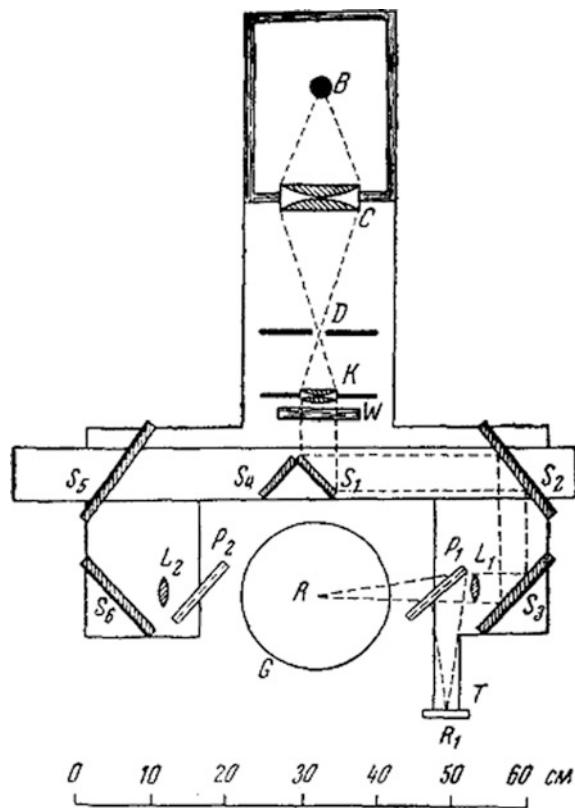
²⁵See W. Crooks, Philos. Transact. of the R. S. of London **170**, 113 (1879).

²⁶See W. Crooks, Philos. Transact. of the R. S. of London **168**, 266 (1878).

²⁷See W. Crooks, cit. above, page 300.

²⁸In my experiments the odds of temperatures between irradiated vane and walls of the bulb were many times more, than the odds of temperatures between two surfaces of the vane. To what function of the first odds of temperatures there corresponds quantity of radiometric forces, their ponderomotive impact on a vane represents their difference on two surfaces of a vane, and this last, with a sufficient degree of approximation, is directly proportional to the second odds of temperatures.

Fig. 12.6 The scheme of the main setup (plan)



the deviation called by a light bundle if the thickness of a vane is equal to zero, that corresponds also to radiometric forces equal to zero. I shall allow myself to note here, that it is necessary to do this corrections only for platinized vanes; at vanes with reflecting surfaces the radiometric forces are so small, against expectation, that they disappear in inevitable errors of observations stipulated by other reasons.

Apart from secondary forces of the known nature mentioned above it is possible to specify also a probable hypothesis, that the pulverization of irradiated bodies, unclosed by the Lenard and Wolf,²⁹ can be accompanied by noticeable reactionary forces, which are inevitable satellites of Maxwell-Bartoli forces of a light pressure; these hypothetical additional forces should, however, depend both on a wave length of an impinging light, and on the chemical nature of a vane; experiments with colour light filters and with different vanes mentioned below have not given opportunities to detect some noticeable impact of these hypothetical reactionary forces.

The general arrangement of devices was the following (Fig. 12.6, plan):

²⁹Ph. Lenard and M. Wolf, Wied. Ann. 37, 455 (1889).

The image of a carbon crater $B(+)$ of the arc lamp (30 A) was projected through the condenser C onto a metal diaphragm D ($d = 4$ mm). The divergent bundle of rays, emergent from a diaphragm, impinged on a lens K and went further as a parallel bundle; to select the bundle from infrared beams there was a glass cuvet, behind a lens K , with parallel plate walls W , filled with pure water³⁰ (thickness of a stratum was 1 cm); to change colouring of beams, it was possible to position in this place additional red (“photographic”) glass or to exchange pure water by a blue ammoniac solution of the copper salt.³¹

On the further path the parallel beam underwent three-multiple reflection from glass (amalgamated) mirrors S_1 , S_2 and S_3 and, being collected through a lens L_1 , gave a real enlarged ($d' = 10$ mm) image R of diaphragms D inside a glass bulb. In the movement of a pair of mirrors S_1S_4 the bundle of rays traveled a similar path and impinged on the other side on a vane located in a glass bulb. The lenses L_1 and L_2 had a focal distance equal to 20 cm each and a size equal to 5 cm; thus a conical bundle of rays had an angle of convergence equal to 15° . All the gadget with mirrors was firmly connected to a lantern of an arc lamp; this last positioned on slides, through which it was easy for removing from a bulb; the adjusting screws and movement on slides allowed to direct a bundle of rays on an explored vane.

It was possible to guard results of observations from influence of those casual springs in luminosity of light, which are inevitably inherent to a volt arc, only by increasing the number of observations.

To refer a separate series of observations to some mean luminosity of light, the following gadget served: between a lens L_1 (Fig. 12.6) and glass bulb the thin flat plate P_1 was posed at 45° to a direction of impinging beams. The majority of light freely transits through a plate; the reflected part of light, being collected, gives a real image R_1 of the diaphragm, which impinges on a thermopile.

The thermopile (Fig. 12.7) consisted of five elements—“constantan-iron” (thickness of wires = 0.025 mm), which were fixed in an ebonite frame and were covered by glass plates; relative luminosity of an impinging light was measured by deviations of the D’Arsonval galvanometer. To attenuate in the same degree the bundle of rays traveling through a lens L_2 , the same glass plate P_2 was inserted here. The luminosity of light was checked only in the case when the pair of mirrors S_1S_4 was in the indicated position (Fig. 12.6); when pair of mirrors was shifted the light could not impinge on a thermopile, and this position served for the definition of a zero point of the galvanometer.

For experiments three different devices (Fig. 12.8) with different vanes were used.

Device I (Fig. 12.8, I) consisted of a glass rod G , to which two crosses made of a leaf platinum of different thickness were pressed by platinum rings (without the

³⁰This procedure eliminated all beams $\lambda > 1.2 \mu$; the glass lenses impede ultraviolet beams at another side.

³¹At red, and also at blue light filter the amount of transiting light energy is reduced up to the one fifth of the white light; It serves the proof that the beams, which were necessary to experiment, almost exclusively belonged to a visual part of a spectrum.

Fig. 12.7 The view of thermopile

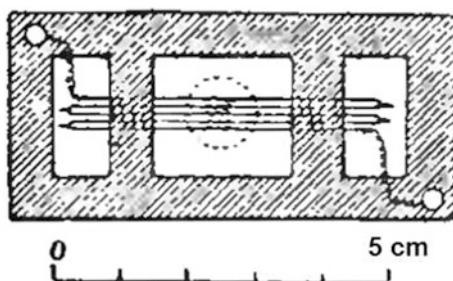
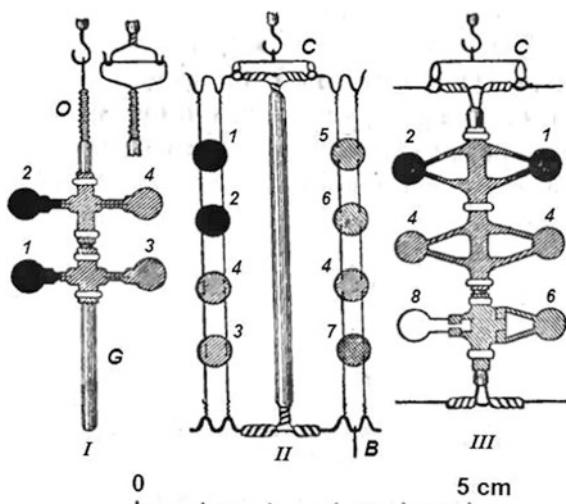


Fig. 12.8 Three types of vanes and supports used in measurements



help of a putty); to make vanes (with diameter = 5 mm) of all the devices isometric, they should be excised by a steel punch. Two vanes of the device *I* had reflecting surfaces from both sides, two others were galvanically covered by platinum niello from both sides,³² whereas the thicker vane exposed five times longer platinization. To suspend the device to a hook of a fiber, the platinum loop *O* was soldered to a glass rod *G*. The loop laid in a plane, perpendicular to a plane of vanes, so that at suspension the rod *G* was erected in a plane of vanes completely freely.

Device II (Fig. 12.8, *II*) also consisted of a glass rod, to which ends the cross platinum wires were soldered. Thin (0.05 mm) platinum wires were tensioned between these holders, which transited through small holes in metal vanes and

³²See F. Kurlbaum, Wied. Ann. **67**, 848 (1899). At the beginning of a platinization it is useful within 30 s to move a vane continuously and strongly in a bath; the surface of a vane gains feeble, grey colouring, like steel. After that the cellular platinum, at a fixed bath, lies on a surface of a vane very strongly.

retained vanes in a vertical plane; these wires were so thin, that their radiometric impacts can be neglected. The device *II* was supplied with a gimbal *C* from a platinum wire, through which it was suspended to a hook of a fiber; the additional platinum bob *B* retained a glass rod in a vertical position.

The device *III* was constructed, as the device *I*, with the only difference, that it was supplied with a gimbal. Narrow metal strips (width 0.3 mm) supporting round vanes ensured a vertical standing of the last in a sufficient measure. The mica vane (8) was inserted into a light casing made of aluminium. The cross wires made of aluminium were attached to a glass rod above and below, so that at omitting the device into a bulb the vanes could not hit about walls of a glass throat.

The experiments were yielded with the following vanes:

Nº	Material	Thickness
1	Platinum platinized by a thick stratum	
2	Platinum platinized five times more thin	
3	Platinum metallic (mirror surface)	0.10 mm
4	Platinum	0.02 mm
5	Aluminium	0.10 mm
6	Aluminium	0.02 mm
7	Nickel	0.02 mm
8	Mica	Thickness < 0.01 mm

The *glass fiber* (length 30 cm) served as a torsion balance which on the low end carried a flat mirror and a hook for suspension of devices. And from the upper side it was fixed in an iron hold-down (Fig. 12.9) inside a mercury section.³³ To attach a fiber *without the help of a putty*, its ends were fixed between slices of an inciderated asbestos board, and these last were pressed below by a platinum ring to the holder of a mirror, and above were seized by a hold-down.

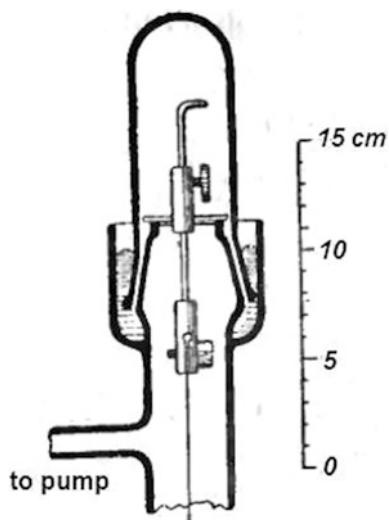
The mirror was positioned in a platinized aluminium casing; it was covered (through an evaporation of the cathode in vacuum) with a stratum of metal platinum, as the silver mirrors were soon attacked by mercury vapours. At a rather weak reflection ability of such a mirror and imperfectionness of the image, due to double passage of a beam through walls of a bulb, an illumination of the scale by Vellman-Martens method³⁴ occurred wonderfully convenient.

The copper wire of length 4 cm was superimposed on a hook of a torsion balance, which mass was equal to 0.314 g in order to determine a value of stiffness from oscillations.

³³All glass sections, executed irreproachably, were supplied by the firm of C. Kramer in Freiburg (Baden).

³⁴F. Martens, Wied. Ann. **62**, 206 (1897); **64**, 625 (1898). The device was obtained from Schmidt und Haensch, Berlin, the price was about 70 marks. I very much recommend a similar scale for operations with sensitive galvanometers and small mirrors.

Fig. 12.9 Scheme of the upper part of the glass bulb with torsion balance



The observations were made with three different torsion balance. The stiffness were so selected, that at distance equal to 1200 divisions of a scale from the scale up to the mirror the double deviation reached from 40 up to 90 divisions of a scale when vanes with reflecting surfaces were enlightened. Thus the periods of one oscillation (in one direction) for the three devices described above were 15, 35 and 13 s.

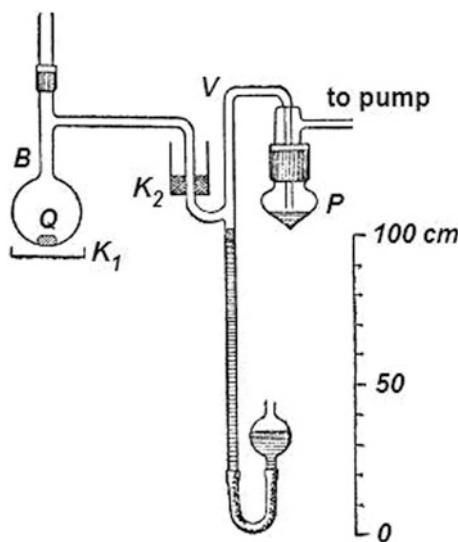
The rarefaction was yielded by the self-acting Kahlbaum pump.³⁵ The pressure measurements made by Leod-Kahlbaum method³⁶ have shown that the rarefactions are easily achieved at which the partial pressure of air is less than 0.0001 mm, (i.e. it is less than the one fifteenth part of saturated mercury vapours pressure at a room temperature).

To receive even greater rarefaction the following trick was used (Fig. 12.10): the drop of mercury Q was located on the bottom of a glass bulb B , then the air was rarefied by the pump, and the mercury drop was heated in water bath K_1 by 5 °C above the room temperature. Being vaporized, the quicksilver is overtaken into the pump and carries away with itself the rest of air from the bulb. If to separate the bulb from the pump and dehumidifier P by a pressure lock V , the ultraviolet

³⁵G. Kahlbaum, Wied. Ann. **53**, 109 (1894). To avoid vapours of lubrication from the cock, which served for a preliminary pumping-out, a barometric lock was arranged between this cock and the pump. An iron parenthesizing into the channel for impinging quicksilver was served as a very essential adding, in views of strength of the pump. The device was obtained from C. Kramer in Freiburg in Br. (Germany). The price was about 350 marks. Being grounded on long-term experiment of operating with self-acting mercury pumps of different types, I should recognize the Kahlbaum pump as the most perfect device of the all known to me, both in care simplicity and in height of achievable rarefaction.

³⁶G. Kahlbaum, Zeitsch. für Instrkd. 15, 192 (1895).

Fig. 12.10 Setup for additional pump of vacuum volume



vapours will stay in a bulb only: their pressure will decrease up to a rather small value if to charge vessels K_1 and K_2 with a cooling intermixture of ice and salt.

The energy of beams, impinging on a vane, was measured calorimetrically: the lantern with mirrors (Fig. 12.6) was removed on slides from a bulb, so that the vane of the device could be substituted by equal D (Figs. 12.11 and 12.12) ($d = 5$ mm). All beams passing through a diaphragm, were absorbed by a calorimeter. The glass plate G compensated decreasing of a light in reflection from a glass wall of a bulb. It was put between a diaphragm and a calorimeter to impede thermal radiation of a diaphragm.

Calorimeter I (Fig. 12.11) consisted of a piece of copper, in which the vertical channel charged with quicksilver was drilled. The blob of the very small calorimetric thermometer, divided into fifths of a degree, was positioned in quicksilver. The absorbing surface of a calorimeter was smoked. The calculated total calorimeter capacity of the device (assuming specific heat capacity of copper = 0.093) was equaled to 3.13 g of water.

Calorimeter II (Fig. 12.12) consisted from the copper cylinder, as well as the first calorimeter, with total thermal capacity equal to 3.61 g of water; its absorbing surface was beforehand gilt, and then it was galvanically covered by platinum niello; this cylinder was put into a copper tube located inside a water bath, about one liter in volume; the bath was supplied with an agitator R . To cool the calorimeter below the bath temperature, prior to begin the experiment, some drops of an etil ether were inlet through a glass tube A into a conical dimple of a calorimeter and then air was blown off by rubber bulb B and banished which carried away the vapour of a volatilizing ether.

The measurements have shown, that from 1.2 up to 1.8 g-cal impinges in a minute on the diaphragm ($d = 5$ mm), i.e. that in my experiments the luminosity of

Fig. 12.11 View of the calorimeter I

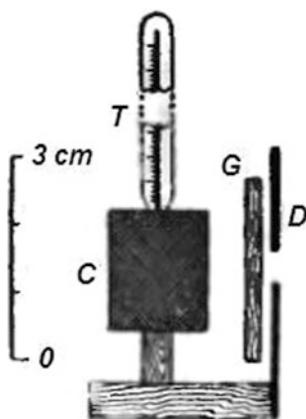
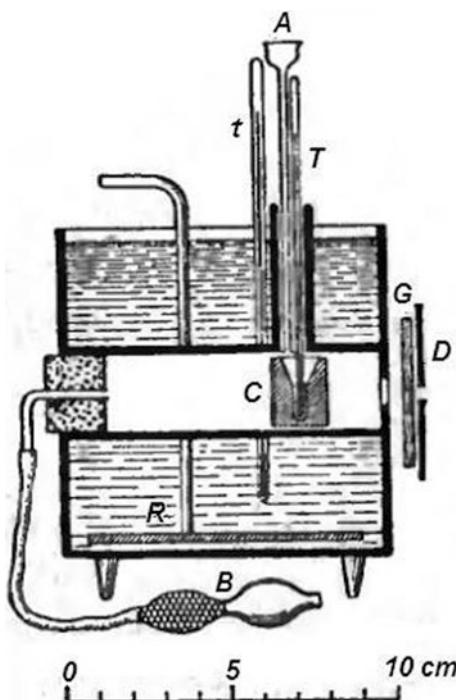


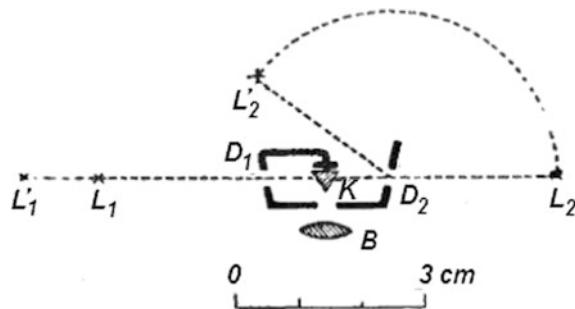
Fig. 12.12 Scheme of the calorimeter II



irradiating was from two to three times higher than the energy of solar beams at a ground surface.

To determine the reflectivity of explored metals the Ritchie photometer (Fig. 12.13) served. The light from two small incandescent lamps L_1 and L_2 impinged, transiting diaphragms D_1 and D_2 (diameter = 3 mm), onto a small prism K made of chalk, and the edge of the last one was observed by lens B . Moving a

Fig. 12.13 Scheme of Ritchie photometer used for reflectivity measurements



lamp L_1 , it was possible to set an identical luminosity. Moving then lamp L_2 approximately by 130° to L_2' and moving up outside an explored metal plate closely to a diaphragm D_2 , it was possible by movement of lamp L_1 in L_1' to set again an identical luminosity. For the angle of incidence equal to 25° the reflectivity was equal to $r = (L_1 K : L_1' K)^2$.

III. Experiments

The gadgets described above allow to solve two basic problems experimentally:

- (1) whether light beams yield any ponderomotive impact independent of the already known secondary forces (convectional and radiometric), and
- (2) whether these new forces of a light meet Maxwell-Bartoli forces of a radiant energy pressure.

Before the beginning of experiments the basic properties of all optical arrangement were investigated preliminary: by moving an additional thermoelement joint with the D'Arsonval galvanometer along the optical axis of lenses L_1 and L_2 (the Fig. 12.6) it was possible to determine their focal distance for the brightest beams of a bundle. Then the mirrors and lenses of the device were positioned so that the real images of a diaphragm on a radiation path both from the right, and from the left quite coincided.

To compare luminosities of beams going from the right and from the left, the additional thermoelement was positioned in the place of formation of real images of the diaphragm. It was alternatively illuminated on the right and on the left. From a large number of measurements it was followed usually, that there is some small odds (about 1%) between the luminosities of both beams. For a large number of reflecting glass surfaces such odds already were due to asymmetrical dust cleaning.

When moving an additional thermoelement by ± 0.5 cm from its main position in the direction of the axis of a beam, in those limits, in which the installations of a beam on a vane could be varied, the diminution of luminosity by 5% was observed for both directions of irradiating.

These preliminary trials were absolutely necessary.

The devices with vanes were always so located inside a bulb, that the beams of a radiant source which missed the vane, reflected and again assembled by a concave wall of a bulb, did not impinge on parts of the suspended device.

After the device with vanes was positioned into a bulb, the pumping out began, proceeding some days, and last pumpings out were yielded at warming up of a bulb walls and at simultaneous irradiating of separate vanes by a light of an arc. Before each series of observations the lower part of a bulb, where there was a drop of quicksilver, was heated in water bath by 5 °C above the room temperature,³⁷ then during from one till two hours the pumping out was again yielded, then the pressure lock V rose, and the cooling by chipped ice and salt followed.

In providing measurements the most essential noises were convectional currents; they have an effect in a continuous drift of zero, and both speed, and direction of this drift depended on casual conditions (even for the same vane per different days of observation). During one series of observations indicated drift of zero happened ordinarily so inappreciable, that, incrementing number of separate observations, it was easy for eliminating. This convection of the residue of mercury vapours was stipulated by heating up of an illuminated vane, and also by casual exterior nonuniform heating up of walls of a bulb and in particular by inevitable odds of temperatures of two cooled mercury surfaces. At observations without cooling the oscillations caused by a convection had an effect much more pronounced, than at cooling by ice with salt; at higher air pressures the observations were so inconvenient, due to a convection, that the measurements were hardly possible.

Another reason causing oscillations of readout was the instability of a voltaic arc, which had an effect even for the best carbons.³⁸ The jumps in luminosity of an arc had an effect in changes (magnification or diminution) of some vibration amplitudes of the device; they were possible for eliminating only by magnification of number of separate observations.

By means of two telescopes the observer could alternately digitize deviations of both devices with vanes and the galvanometer. An assistant,³⁹ observing for correct burning of an arc, translocated a pair of mirrors S_1S_4 (Fig. 12.6) on a command. Making irradiatings with periodic interruptions, it was possible to amplify a vibration amplitude of the device to the necessary value.

The Table 12.1 shows a beginning of one of the protocols of observations.

Notations of this table are:

L_1 and L_2 are turning points on a scale, when the light have impinged on a vane of the device from a lens L_1 or from a lens L_2 . A middle column, “*calculations*”,

³⁷At the indicated small odds of temperatures quicksilver is not precipitated on more cold walls of the device; this appearance, having place at unwettable surfaces, was indicated by M. Cantor (Wied. Ann. **56**, 493 (1895)).

³⁸Quite satisfactory there were Simmens carbons “A”; with cheaper carbons observations are hardly possible.

³⁹My assistant at these experiments was the preparator assistant Avtonom Fedorov; his diligent attitude and dexterous treatment with devices was appreciably facilitated to me these uneasy observations.

shows the equilibrium point calculated (from three adjacent tuning points). “*Deviation*” means a deviation of system at opposite direction of irradiation.

G_1 and G_2 give readouts of the galvanometer in the first and second cases (in the second case it is the origin).

“Galvanometer” give deviations of the galvanometer.

“Deviations relative ($G = 100$)” give the above deviations of the device, relative to a fixed deviation of the galvanometer of 100 divisions of a scale.

By the method indicated in Table 12.1 it was performed seven ordinary readouts for L_1 and L_2 and the mean value was derived from “*Deviations relative ($G = 100$)*” with a mean \pm deviations of individual observations. (For a vane of the Table 12.1 this double deviation was $a = 29.4 \pm 1.6$ of scale divisions.)

Table 12.1 Data on angular deviations of torsion balance, divisions of a scale

Device III. A platinized vane (2). Distance of centre of a vane from a rotation axis = 9,2 mm. Cooling by ice with salt.

Distance up to a scale A = 1195 divisions of a scale.

L_1	L_2	L_1	L_2	L_1
calc.	calc.	calc.	calc.	.
306	115	206	295	307
176	240	184	245	174
239	302	118	207	210
177	239	208	296	244
240	302	124	209	211
178		294	189	212
240		208		245
240		244		213
240				247
240				.
Deviation: 32 divisions		36 divisions		32 divisions
G_1	G_2	G_1	G_1	G_2
308		314		.
305	201	312	201	.
312		314		.
314		316		.
310		314		.
Galvanometer: 109 divisions		113 divisions		113 divisions
Deviation relative:				
(G - 100) 29,3 divisions		31,8 divisions		28,2 divisions

To compare observations made with different vanes the following additional corrections were necessary.

In Devices *I* and *III* the narrow beam of light impinges, apart from the circle, on the parts, supporting it, due to that the deviation is incremented; by measuring the areas of enlighten parts and their distance from a rotation axis we can subtract that additional impact, which they yield (from 5% up to 10% of the total quantity), and we gain that deviation, which is caused by a circle of a vane only (device *II* is free from this correction). For a vane of Table 12.1 this correction makes 1.9 divisions of scale; the calculated double deviation is 27.5 divisions of scale.

The measurement of distances from the circle center of a vane to a rotation axis was yielded by the following way: the arc lantern with the reflecting device was removed on slides off, and from the side of beams, impinging during experiment, the plumb-line made of a thin bright silver wire was hung up as close as possible to the bulb; the telescope with an ocular micrometer was placed perpendicularly to a plane of disks at a distance about 4 m, and it was necessary to move a plumb-line until then, it covered a fiber. The quantity relevant to one division of an ocular micrometer of a telescope, was determined with the help of sighting a scale located near a bulb; a visible distance of a circle center of a vane from a plumb-line gave true distance from the first one to the torsion axis and could be measured to within ± 0.5 mm; the measured distances laid between 9 and 11 mm.

On the basis of these measurements the observed double deviations were given in deviations relative to distance of centers of circles from a rotation axis, equal to 1 cm. For a vane of Table 12.1 such relative deviation was equal to 29.9 divisions of scale.

To determine an absolute value of light pressure on a vane, it was necessary to measure an absolute value of the rotational stiffness of fiber. In place of the device with vanes a solid body (copper cylinder) with a known moment of inertia was suspended to a hook of a torsion balance, and from three series of observations, of which everyone consisted of ten oscillation periods, the mean time of one period was derived⁴⁰ (Table 12.2).

On the basis of the indicated value of rotational stiffness we gain for a vane of Table 12.1 under *unilateral* irradiating the value of a light pressure in dynes:

$$p = 0.0000308 \text{ Dynes} \pm 0.0000017 \text{ Dynes}$$

To test calculations of Maxwell and Bartoli, it is necessary to estimate the value of light pressure, which would be expected in experiments according to the mentioned theory, and to compare the calculated value with the observed one. For this purpose it is necessary to make a calorimeter measurement of impinging light energy, and also a photometer measurement of reflectivities of the vanes.

⁴⁰Compare: F. Kohlrausch, Lehrbuch der praktischen Physik, § 29 and comment 11 and 12, Teubner, Leipzig 1901.

Table 12.2 Data on oscillation period and rotational stiffness of the fiber

Oscillation period	
Only mirror $\frac{t_1}{2} = 5.1 \pm 0.05$ s	Copper cylinder
Mirror + copper cylinder $\frac{t_2}{2} = 29.4 \pm 0.1$ s	Length = 4.0 cm Mass = 0.314 g
Rotational stiffness $D = 0.00494$ Dynes cm	

The measurements made with the help of the first calorimeter (Fig. 12.11) were yielded as follows: the mirrors (Fig. 12.6) were moved aside on slides so, that it was possible to put a diaphragm of a calorimeter D in the place of devices with vanes. Then the calorimeter was illuminated within 5 min, and every minute the observations of the thermometer (together with galvanometer) were made. After that the irradiating was interrupted by means of the opaque screen, and in the following 5 min the observations of the thermometer, which now gradually diminished, were made every minute again, and the zero point of a galvanometer was observed. A complete series of observations implied five sequential periods of irradiating.

All observations were handled pictorially, for that purpose the observations of the thermometer were plotted on a coordinate paper and were joined by a continuous curve so that the last one flowed as smooth as possible (Fig. 12.14). It is clear from the figure that the course of temperature in 10 s exhibits a transition from irradiating to a blackout or back by a specific turning point.

The very high velocity of a calorimeter cooling requires the special treatment of results, as even during one interval of observation neither velocity of heating up, nor velocity of cooling are not constant values. For a definite mean temperature of a surface of a calorimeter both velocities have constant values represented by tangential lines (the last ones are easily superimposed on the drawing). For these constant values the intersection points of tangential lines with ordinates, restricting the interval, gave those temperature differences, which would be established in 5 min, if both velocities were constant. The sum of two differences gives a total, corrected by losses, rise of a calorimeter temperature.

But here a source of errors in determination of true mean temperature of a surface appears; the thermometer has not enough time to follow the temperature and it gives at irradiating too low, and at cooling too high observations. That circumstance, that the thermometer follows a turning point in 10 s, allows, as a first and for our experiments a sufficient approximation, to suppose, that the thermometer lags behind on 20 s. Then for the mean temperature it is necessary to compare not the points of a curve t_1 and t_2 , but points T_1 and T_2 , laying on the same curve by 20 s earlier.

Such pictorial treatments were done at each heating up for two temperatures; Table 12.3 represents one series of measurements.

With the second calorimeter (Fig. 12.12) the measurements were much easier: the calorimeter was cooled (by 2.5° below the bath temperature) with the help of an ethyl ether, then exposed to heating up by beams, and the observer made the readout of the calorimetric thermometer each minute (and in gaps—a deviation of

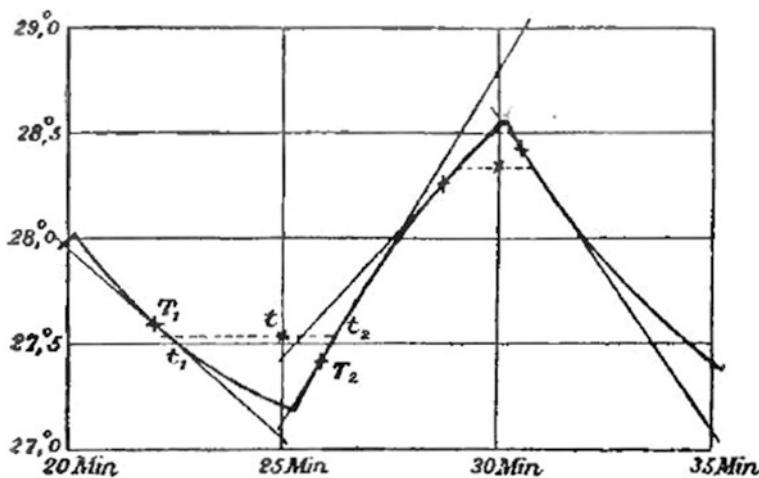


Fig. 12.14 Diagram of time-dependent thermometer data and their treatment

Table 12.3 Results of thermal measurements of light power by calorimeter I

Calorimeter I. Water equivalent = 3,13 gr.

	Speed of heating	Speed of cooling	Total heating	Deviation of galvanometer	Relative total heating (G = 100)
I {	1 ^o ,57	0 ^o ,63	2 ^o ,30	140 дел.	1 ^o ,64
	1,49	0,80	2,29	128 "	1,79
II {	1,44	0,85	2,29	128 "	1,79
	1,31	1,10	2,40	122 "	1,97
III {	1,38	1,08	2,46	129 "	1,91
	1,00	1,37	2,37	126 "	1,88
IV {	1,30	1,15	2,45	123 "	1,99
	1,04	1,45	2,49	127 "	1,96
V {	1,26	1,27	2,54	129 "	1,97
	0,93	1,50	2,43	126 "	1,93

Total heating in 5 min (G = 100) $1^o,88 \pm 0^o,09$

Table 12.4 Results of thermal measurements of light power by calorimeter II

Calorimeter II. Total calorimeter capacity = 3,61 gr. of water

Series of observations	Heating up in 5 min.		Mean	Deviation of galvanometer	Heating up relative to G = 100 divisions of scale
	from tangential line	from differences of temperatures			
I	2°,40	2°,41	2°,40	159 дел. скал.	1°,51
II	2°,55	2°,57	2°,57	163 " "	1°,57
III	2°,43	2°,50	2°,46	158 " "	1°,56

Mean heating up (for $G = 100$) $1^{\circ},55 \pm 0^{\circ},02$

the galvanometer and the temperature of a water bath). The observations were plotted, joined by a continuous curve; on this curve the bath temperature was fixed,⁴¹ and at this point a tangential line was carried out to a curve relevant to the *true* velocity of heating up of a calorimeter, irrespective of losses through a radiation. If to take two points of a curve relevant to time 2.5 min before and later the temperature equalizing, we also receive a mean velocity of heating up of a calorimeter during 5 min. Table 12.4 gives the results of observations.⁴²

From here we obtain the value of energy, impinging within second:

$$E = \frac{1^{\circ},55 \times 3,61 \times 4,18 \times 10^7}{300} \text{ Ergs} = 7,74 \cdot 10^5 \text{ Ergs.}$$

At our experiments the beams impinged not as collimated, but as a convergent one; their declination was, however, so inappreciable, that a correction caused by it⁴³ (about 1%) could be dropped in view of other much larger inaccuracies of observations. We can, hence, make calculations according to the formulas, given by Maxwell and Bartoli for a collimated beam.

⁴¹Again it is necessary to have in mind that the calorimeter thermometer is in delay from the true temperature of a calorimeter by 20 s.

⁴²Results of Tables 12.3 and 12.4 can not serve for immediate comparison, as they concern to different adjustments of a thermopile.

⁴³See L. Boltzmann, Wied. Ann. 22, 292 (1884), and also D. Holdhammer, cit. above, page 844.

Table 12.5 Data on reflectivity of vanes

	Photometer measurements		By Hagen and Rubens	
	ρ	ρ (MB)	ρ	ρ (MB)
Platinum	0.5 ± 0.05	1.5	0.64	1.64
Aluminium	0.6 ± 0.05	1.6	0.83	1.83
Nickel	0.35 ± 0.05	1.4	0.65	1.65

For the absolute black body we gain on the basis of calorimeter measurements of Table 12.4 a value of pressure p :

$$p \text{ (Dynes)} = \frac{E \text{ (Ergs)}}{3 \cdot 10^{10}} = 0.0000258 \text{ Dynes.}$$

To express the results obtained in conveniently **comparable** quantities, we shall take as a *unit* of comparison the value of Maxwell-Bartoli pressure referred to the *absolute black body*, calculated from calorimeter observations, and we shall term this interim unit as MB unit.

In these units the results of Table 12.1 will be expressed as follows:

$$p = \frac{0.0000308 \pm 0.0000017}{0.0000258} = (1.19 \pm 0.07) \text{ MB.}$$

The straightforward measurement of reflectivities of explored vanes was impossible, because their surfaces have appeared too rough. Therefore I have measured by a photometer (Fig. 12.13) reflectivities of those metal leafs, of which the vanes were made; irregularities of these leafs also had a substantial effect together with clearly observable colouring of a reflected light (especially for nickel); the values of these reflectivities measured for an angle of incidence 25° , are given in Table 12.5 without further reductions. For the comparison reflectivities here are also indicated for a normal slope of beams ($\lambda = 600 \mu\mu$) according to Hagen and Rubens,⁴⁴ and on their basis the Maxwell-Bartoli forces are calculated (values obtained for magnalium are given for aluminium).

I do not give evaluations for *mica*, as the observations were made only with one vane, and there are no test measurements with thicker vanes.

The results of a individual series of observations made by me with different devices are presented below. When I had transferred from observations at room temperature, at which the inevitable oscillations of final outputs are rather significant, to measurements with cooling by ice with salt, I did not expect to receive such

⁴⁴H. Hagen and Rubens, Ann. d. Phys. 1, 373 (1900).

agreement between the observed quantities and those calculated according to Maxwell-Bartoli, which resulted from my experiments; I therefore have assumed, that such coincidence of evaluations and observations is necessary to assign to accidents, and consequently at first has exchanged *I* calorimeter by *II* calorimeter, and then *II* device with vanes by *III* device.

The numerous observations, which I made with *I* device at a room temperature, were not so good as the subsequent measurements, and therefore they were not given by me here. The observations with a platinized vane (2) of *II* device were not given also, as at the microscopic examination of a vane, which had followed the experiments, it was found, that the platinum niello had subsided unsatisfactorily as a sponge (that was not observed on other vanes). With *III* device, unfortunately, only two series of observations were made, as the further experiments were interrupted by breakage of the device.

IV. The Results

The results of experiments are given in terms of MB units; the mean deviation of devices is given in the same units under every observed quantity, whereas all deviations, smaller than 0.15 MB are designated as 0.1 MB; those below 0.25 MB are designated as 0.2 MB and so on.

The following reasons could serve to get an idea about the precision of the given measurements: the deviations at installations of the device during measurements were given in Table below; the determination of an absolute value of a pressing force of light (where measurements of the rotational stiffness of a fiber enter, and measurements of the distances from a mirror up to a scale and the distance from the center of a vane up to a rotation axis) was possible to be made with precision about $\pm 8\%$; evaluation of an absolute value of MB unit from calorimeter measurements (which include a total water capacity, the increase of temperature of a calorimeter and the ratio of the area of a diaphragm to the area of a circle of a vane, which was close to unity) was possible to be made with probable precision in $\pm 7\%$; the inaccuracy in definition of true value of reflectivities, probably, did not exceed $\pm 10\%$.

Random inaccuracies of settings of the real image center of the diaphragm onto the vane were added to the indicated inaccuracy of individual measurements and also the possibility, that the radiation reflected from a vane was focused by a concave surface of a bulb on other parts of the suspended device, and the point of this secondary heating up varied during one oscillation of the device. The total random error, at the described measurements with *a white light*, probably, did not exceed $\pm 20\%$.

		Device I	Device II	Calorimeter II			Device III		
		Calorimeter I	White light	White 1.	Red light	White 1.	Blue 1.	White light	
1	Thick-platinized vane	1.8 ± 0.2	6.1 ± 0.1	1.5 ± 0.1	—	—	—	1.5 ± 0.1	1.4 ± 0.1
2	Thin-platinized vane	1.3 ± 0.2	—	—	—	—	—	1.2 ± 0.1	1.1 ± 0.1
Calculated		1.2	—	—	—	—	—	1.1	1.0
3	Platinum, thick	—	1.8 ± 0.1	—	—	—	—	—	—
4	Platinum, thin	—	2.0 ± 0.1	1.9 ± 0.2	18 ± 0.1	1.9 ± 0.1	(1.8 ± 0.8)	1.7 ± 0.1	(1.5 ± 0.5)
5	Aluminium, thick	—	2.3 ± 0.4	1.9 ± 0.1	—	—	—	—	—
6	Aluminium, thin	—	2.0 ± 0.1	2.3 ± 0.1	2.0 ± 0.2	(2.9 ± 0.8)	2.1 ± 0.1	(2.5 ± 0.5)	1.4 ± 0.2
7	Nickel, thin	—	1.7 ± 0.3	1.2 ± 0.2	1.4 ± 0.1	(2.3 ± 0.5)	1.4 ± 0.2	(2.7 ± 0.9)	—
8	Mica	—	—	—	—	—	—	0.08 ± 0.05	0.13 ± 0.03

In experiments with red and blue light, when the amount of impinging energy was five times less, casual oscillations caused by a convection, were the same, and consequently, the precision of the obtained results was correspondingly less; it was necessary to note the same also on rather very small deviations (hardly reaching four divisions of a scale) at a mica vane. These experiments, which were undertaken as test ones, nevertheless allowed to state, that in these cases there were no new ponderomotive forces which would be comparable to the Maxwell-Bartoli forces in their value.

Besides, I multiply provided comparative measurements with thin and thick metal (reflecting) platinum and aluminium vanes; however, I did not manage to detect clearly enough expressed radiometric odds; that was why it was possible to consider radiometric forces of thin metal vanes as equal to zero within limits of observational errors. The results obtained can be stated as follows:

- (1) The impinging beam of light yields pressure both on reflecting, and on absorbing surfaces; these ponderomotive forces are not related with already known secondary convectional and radiometric forces caused by heating up.
- (2) The forces of light pressure are directly proportional to the energy of an impinging beam and do not depend on its colour.
- (3) The observed forces of light pressure, within limits of observational errors, are quantitatively equal to the Maxwell-Bartoli forces of pressure of a radiant energy.

Thus the existence of the Maxwell-Bartoli forces of pressure has been established for the light beams experimentally.

Physical laboratory of the University. Moscow, August 1901.

Chapter 13

First Experiments on Measuring Light Pressure II (Ernest Fox Nichols and Gordon Ferrie Hull)



Elsa Garmire

Abstract This chapter contains an essay of Elsa Garmire about Nichols (1869–1924) and Hull’s (1870–1956) results on measurements of light pressure in 1901 and later and their comparison with Lebedev’s results (Sect. 13.1). It also contains a reprint of the paper of Nichols and Hull of 1901 in which their first results were reported (Sect. 13.2).

13.1 Musings on the First Measurements of Light Pressure

13.1.1 Introduction

The saga of the first successful experiments on light pressure frame interesting issues that I invite you to ponder. James Clerk Maxwell theoretically proved in 1873 that transverse light waves exert pressure proportional to their energy density. By 1900 the existence of light pressure was well accepted, but calculations showed that light pressure would be very small and numerous experiments had failed to observe it. This was a great challenge. How could experimenters be sure they had observed it? Would numerical agreement with Maxwell’s predictions provide greater “proof” than pure observation? How much accuracy would provide “proof”? Twenty percent? One percent? Who should get credit for its discovery? The one who first measured anything, or those who made the first accurate measurements? What if the paper claiming the first “accurate” measurements contained errors?

E. Garmire (✉)

Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA
e-mail: elsa.garmire@dartmouth.edu

E. Garmire
1156A East Cliff Drive, Santa Cruz, CA 95062, USA

By 1900 equipment was finally reaching capability of measuring the extraordinarily weak pressures. A Russian and a pair of American researchers were unknowingly competing to be first; approaching their experimental design very differently. Their different “taste” in experiments was based on availability of equipment, personal choice, their local scientific culture, previous experience, political, financial, and other considerations. What approximations and/or assumptions would have to be made? The more complex the experiment, the more complex the data analysis and the greater chance that many silly errors might sway the answers. Would this matter?

The aim was to numerically demonstrate the Maxwell-Bartoli¹ equation (hereafter called MBE), which related the light pressure to the energy density: light pressure incident upon a reflecting plane with reflectivity r is $p = E/V(1 + r)$, where E/V is the light energy density. The MBE gave a numerical target for experimental measurements. Did it matter that MBE had remained experimentally unproven? No one questioned MBE theoretically. What accuracy would be needed to “prove” it?

13.1.2 First Archival Reports of Light Pressure

The first report of successful experiments to measure light pressure and “proof” of MBE was presented in Russia by Peter Lebedev² from the University of Moscow in 1899, with his discovery announced to the rest of the world at the International Congress on Physics in Paris in 1900 [1]. His final results were archivally published in German, November of 1901 [2], having been submitted in August.

Meanwhile, Gordon F. Nichols and Ernest F. Hull (unaware of the Russian work) were making similar measurements at Dartmouth College in the United States, using a substantially different approach. They submitted their work for publication [3] at the same time as Lebedev, also in August, 1901. Their results came out in Physical Review in November of 1901, simultaneously with the Lebedev published results [3], but half a world away. Because of the simultaneity (and independence) of the two publications, both authors and both countries are usually given equal credit for these initial measurements. However, we will find that many years later the paper of Nichols and Hull (called N&H from hereon) was called into question. We’ll have to decide if this should invalidate the credit they received.

¹Bartoli was an Italian who provided a proof of this equation for all waves by means of thermodynamics at roughly the same time as Maxwell proved it by electromagnetic theory.

²Today’s spelling will be used for his name, although at the turn of the century his name was transliterated as *Piotr Lebedew*.

13.1.3 Comparing Experiments

Both experiments relied on the radiometer invented by Sir William Crookes in 1874. Previous investigations had found that pressure from residual gas³ masked the extraordinarily small radiation pressure. They both relied on carbon arc lamps as their source and calorimetry to measure illumination energy, but their experimental techniques were quite different. Comparing these provides valuable lessons on experimental design challenges; Lebedev's measurement techniques were much simpler, while N&H's seemed to be more definitive.

Lebedev Results

To remove gas effects, Lebedev evacuated his flask as far as was technically possible, below 0.0001 mm, using a newly developed mercury diffusion pump in the vacuum line. With such a low flask pressure, he argued that gas convection would be small. He further found he could reduce heating of the vanes by making them thermally conductive and as thin as possible (thin platinum vanes).

Any remaining gas effects were canceled out by ensuring the vanes were as vertical as possible and letting the radiation fall upon each side in turn. He compensated for the illumination fluctuations by introducing a beam splitter and monitoring the energy in real time as he was monitoring pressure by deflection of the vanes. Light energy was determined from the temperature rise determined by thermopile currents.

Lebedev made continuous measurements with his arc lamp, reporting a maximum of 7 mrad deflection. The force required to achieve the measured time-average deflection was compared with the force calculated from MBE. The measured values agreed with MBE within the estimated 20% experimental error. Lebedev ended his paper with "Herein is the existence of Maxwell-Bartoli's equation for the compressive force of light experimentally proven"⁴" [2]. This paper is usually given credit for the first demonstration of light pressure. Extensive details of his experiment were published in 1902 in English in the Chemical News and Journal of Physical Science, edited by Sir William Crookes [4].

Nichols and Hull Results

Nichols and Hull did not have a vacuum system capable of ultra-low pressures; they needed elaborate and ingenious methods to surmount gas effects. They took dual approaches. First, they extensively studied the effect of gas pressure and found

³The Radiometer consisted of four vanes, each of which was blackened on one side and silvered on the other, and attached to the arms of a rotor balanced on a vertical support designed to turn with very little friction. The mechanism was placed inside a clear glass bulb and pumped out to a high, but not perfect, vacuum. Theory shows that in vacuum, radiation pressure would be twice as high on the reflecting side, so the high reflectivity side would be pushed away. In the radiometer, however, residual gas will heat up due to absorption on the black side and push the vanes away.

⁴Garmire's translation from the German.

some pressures where gas pushed the vanes and other pressures where it offered suction; they identified a pressure where the gas effects were minimized. Secondly, they added a shutter to pulse illumination on the vanes and made time-dependent measurements of vane deflection. They evaluated vane motion at short times before thermal effects could build up. To measure light energy they used a home-made platinum bolometer⁵ designed by Professor Nichols, who had already achieved international fame with his earlier work on these sensitive devices. To compensate for unsteady source intensity, they rapidly shifted from radiometer to bolometer and back. Their typical maximum torsion angular deflection was 28 mrad (a 3 cm deflection in a 105 cm long arm); its time dependence made its interpretation much more complex than Lebedev's. In their initial 1901 report they reported results about 20% higher than theory, which they attributed to unknown errors in interpretation of their experiments, but added "The writers believe that the observations already in hand are sufficient to prove experimentally the existence of a pressure, not due to gas molecules, of the nature and order of magnitude of radiation pressure, but toward a close quantitative measurement of this pressure much remains still to be done" [4]. It is interesting to note that they never questioned whether the theory could be wrong!

13.1.4 Toward Numerical Validation of Radiation Pressure

Further Nichols and Hull Results

N&H continued to refine their experiments and analysis, presenting their results at the American Academy of Arts and Sciences [5] in December, 1902, which were published in 1903 [6]. They provided full details of their experiments and justification for their analysis, correcting the error in their first paper (determining the area of their bolometer). Now their experimental results matched MBE within 1%, which they claimed was comparable to their experimental errors. They believed these were the first final and definitive experimental results proving MBE.

In their 1903 paper, N&H gave reference to Lebedev's early paper, but stated: "No estimate of the 'errors of observation' was given in the paper nor other numerical data. Unfortunately the proceedings of the Paris Congress did not reach the writers nor any intimation of the methods or results of Professor Lebedev's work until after the publication of their own preliminary experiments."

⁵The light illuminated a platinum disc allowing it to heat. When the light was turned off, electrical current was passed through the platinum disc until it heated to the same temperature. Knowing the resistance of the disk and the amount of current, the ergs of absorbed illumination could be determined.

The 1903 results were immediately accepted as definitive and widely circulated within the Physics community. In 1904, a paper in the Astrophysical Journal said “Quite independently of Lebedev, Nichols and Hull have been at work upon the same problem, and with such refinement of method that they have just been able to announce perfect agreement with the theoretically established value within the limits of their observation errors (about 1%)” [6, 7]. Such a confirming measurement was prized by astrophysicists who were predicting astronomical behavior. Again, in 1905, J. H. Poynting (well-known from electromagnetic theory) referenced “the great experiments of Lebedev and of Nichols and Hull,” and concluded that a “wave train may be regarded as a stream of momentum travelling through space,” during his Presidential Address before the Annual General Meeting of the Physical Society [8]. The 1903 N&H paper became dogma claiming about 1% accuracy of light pressure measurements, particularly in the US.

Nichols and Hull had achieved great fame for their heroic experiments and accurate results, as did the Dartmouth Physics Department. However ten years later these results began to be called into question. Was the Maxwell-Bartoli equation indeed proven to 1%? The question of whether the Russians or Americans did it first boils down to what is meant by scientific “proof.” Is 20% enough, or 1% better? The true story was confused even more by unwitting errors in both the 1901 and 1903 papers of N&H. It would be shown eventually that their careful but circuitous experimental methods would not have the accuracy they thought. They were overly-optimistic and made other errors that would need correction. It would not be until 1934, however, that their published papers would be questioned and their data re-evaluated.

Golsen’s New Measurements of Light Pressure

Ten years after the publications of Lebedev and N&H, in 1924, improvements in technology inspired Professor Walther Gerlach to repeat the earlier measurements of light pressure, striving for greater accuracy. Vacuum systems could now reach pressures as low as 10^{-7} mm Hg, better lamps were available and Gerlach had designed more sensitive thermopiles. Fraulein Alice Golsen made the measurements following the straight-forward procedures of Lebedev. She and Gerlach reported the following: (1) radiation-induced deflections independent of vacuum pressure from 10^{-6} to 10^{-7} mm Hg; (2) deflection angle was proportional to light energy and independent of wavelength; (3) deflection was due to radiation pressure only and agreed with the theoretical value determined from MBE [9]. Their estimated accuracy was $\pm 2\%$, ten times more accurate than Lebedev had claimed. Golsen detailed her experiments in a separate paper, explaining that she measured standard deviations a bit over 1%, but that possible experimental errors could be as large as 6% [10]. They might have wondered if N&H had over-reached in their claims of accuracy, because they could do no better, even with their advanced

technology. They apparently didn't notice, however, and the N&H claim of 1% accuracy remained as dogma for another ten years.

13.1.5 Mary Bell and S. E. Green Re-analysis of the Nichols and Hull Papers

Things changed, however, in 1933, when Mary Bell and S. E. Green were experimenting with the limits of their own radiometer and analyzing its potential. They pointed out that "While the above preliminary experiments were in progress, the errors of calculation occurring in Nichols and Hull's paper on radiation pressure were noticed. Although their work, carried out at a gas pressure of 16 mm of mercury, has been widely quoted as conclusively establishing, to within about 1%, the numerical equivalence between the pressure and energy-density of radiation, it is found that their results, when correctly evaluated, show a divergence between these quantities of some 10%. Hence Nichols and Hull's investigation cannot be regarded as furnishing a quantitative experimental verification of the equality relationship deducible from theory" [11].

Wow! Bell and Green were questioning beliefs that had been in place for 30 years! Indeed, they had identified four N&H errors: (1) the ballistic equation of motion for the swinging vane was missing a term that was not small; (2) the calculation for the maximum deflection of the initial swing of the vane was incorrect; (3) the calculations incorrectly used log base 10 rather the natural logarithm (base e); (4) The wrong mechanical equivalent of heat was used. When Bell and Green re-evaluated the N&H paper correcting these errors, they found 10% deviation from the Maxwell-Bartoli equation. The N&H claim of 1% error for their measurements had been overly optimistic!

Bell and Green had dropped a bombshell! Particularly on Professor Hull, who was a professor at Dartmouth and whose reputation was at stake. He got together with Bell and Green and in 1934 the three published a combined paper in which each got a say. The abstract of this paper said: "The first part consists of a brief account [by S. E. Green] of some early experiments on radiation pressure, dealing in particular with the investigations of Lebedev and of Nichols and Hull. In the second part, Dr. Hull reconsiders some aspects of his work, in view of the errors of calculation pointed out by Mary Bell and S. E. Green. In the third part the authors reply to Dr. Hull's note" [12]. Green wrote, *inter alia*, "The best known investigation on the subject [light pressure] is that of Nichols and Hull, who used great ingenuity and experimental skill in an attempt to obtain a quantitative test of Maxwell's relation. ... The circuitous experimental methods made the evaluation of

the final results very involved. ... Correct evaluation of Nichols and Hull's results shows the pressure to exceed the energy-density by from 9 to 11% of the latter."

Hull did admit to errors: "As a mathematical performance our solution for the general case was very faulty." He does complain, however, that "Lebedev, whose paper on light pressure appeared about the same time as our first paper, rejected as small, quantities less than 10% and had variations between the two values of light pressure as high as 50% for white light and 80% for blue light. Yet his work is quoted at times as having verified the Maxwell-Bartoli relation." In his conclusion, he justified N&H: "Again, when all corrections have been applied ... the agreement is within 7%. In view of the very much larger variations in corresponding absolute measurements of radiation which prevailed at that time it can be stated that the experiments of Nichols and Hull verified the theoretical relation of Maxwell and of Bartoli."

Mary Bell attempted to mollify Prof. Hull: "Dr. Hull states that the final agreement was fortuitous, and this is obviously so since, all other corrections apart, the new corrections to the bolometer resistance (admitted by Dr. Hull to be a very critical factor) bring into agreement pressure and energy values which otherwise would differ by 20%. ... The errors we pointed out were merely of calculation, and do not detract from the experimental ingenuity with which Nichols and Hull conducted their classical investigation. ... We do, however, claim for our own methods the great advantages of simplicity and directness, with the consequent possibility of avoiding the majority of the complicated corrections inherent to the work of Nichols and Hull."

13.1.6 *Philosophers and Historians Have Their Say*

Science historian G. Burniston Brown pointed out: "There does seem to be a real danger of finding what you expect to find when the effects are very small. ... Six pages of [the Bell and Green] paper are devoted to describing the errors made by Nichols and Hull, which included muddling \log_{10} with \log_e , mA/mm with mm/mA for the galvanometer sensitivity, etc. Hull was forced to admit that the agreement of his results with Maxwell's Theory was 'fortuitous.' ... This experiment is actually a classical example of the tendency to verify the predictions of a theory already believed in. ... Had there been disagreement with theory these errors and omissions would have been discovered and rectified" [13].

Philosopher Morton L. Schagrin found that "The pressure of light seemed to be an ideal subject for investigating questions such as the historical role of theoretical expectations and the impact of observational results on theories" [14]. Philosopher John Worrall wrote: "These later pressure experiments are not, then, to be dismissed

as unimportant for the appraisal of the merits of our present theories of light—they do provide them with some support. Nonetheless the relationship between theory and experiment has turned out to be much more complicated and subtle than might have been expected. Certainly the full analysis of these experiments has taken us a long way from the initial, naive idea that they could operate as ‘crucial experiments’, straightforwardly supplying a criterion of ‘truth’ or of ‘falsity’ for basic theories of the constitution of light” [15].

13.1.7 Conclusions

I believe this story offers numerous insights, beginning with the initial conundrum behind light: is it a particle or a wave? How can it be that light exerts a pressure? Secondly, what does it mean to “prove” a theory? If this is the first demonstration, then how do you know you are measuring the desired effect? What kind of numerical agreement is expected?

Who should get credit? Who does get credit? The first to claim it? How accurate must the first experiment be to claim credit? Should more credit go to the first to make accurate measurements to prove agreement with theory? In fact, 180 papers have cited Lebedev’s paper (ISI statistics), while only half as many have cited N&H’s first 1901 paper (88), and only a third (57) have cited their second 1903 where they quote 1% error. The finally corrected Hull paper in 1934 has been cited only twice! I leave you to draw your own conclusions as to what this means about science.

Do cultural factors influence who gets credit? In 1901 Lebedev’s work was published in German journals, at a time when German science dominated. N&H published in America, a country which was not particularly respected scientifically. They could not get to France to attend Lebedev’s first international talk—nor did they even hear of it. Travel was by boat only. Scientific information also came by boat, taking several months. When Lebedev and N&H simultaneously presented their published results, the German journal was more internationally prestigious and garnered the most attention (and still does today).

Considering the rest of their careers, Nichols took to administration, eventually becoming President of Dartmouth, and then of MIT. Hull spent fifty years teaching physics at Dartmouth. By contrast, Lebedev, as a matter of principle, resigned from his faculty position in 1911, protesting against the politics of the Ministry of Education (according to Wikipedia). This was, of course, in the middle of the Russian revolution. Sadly, he died in 1912, limiting his contributions to Physics primarily to his one remarkable determination of light pressure. However, his name

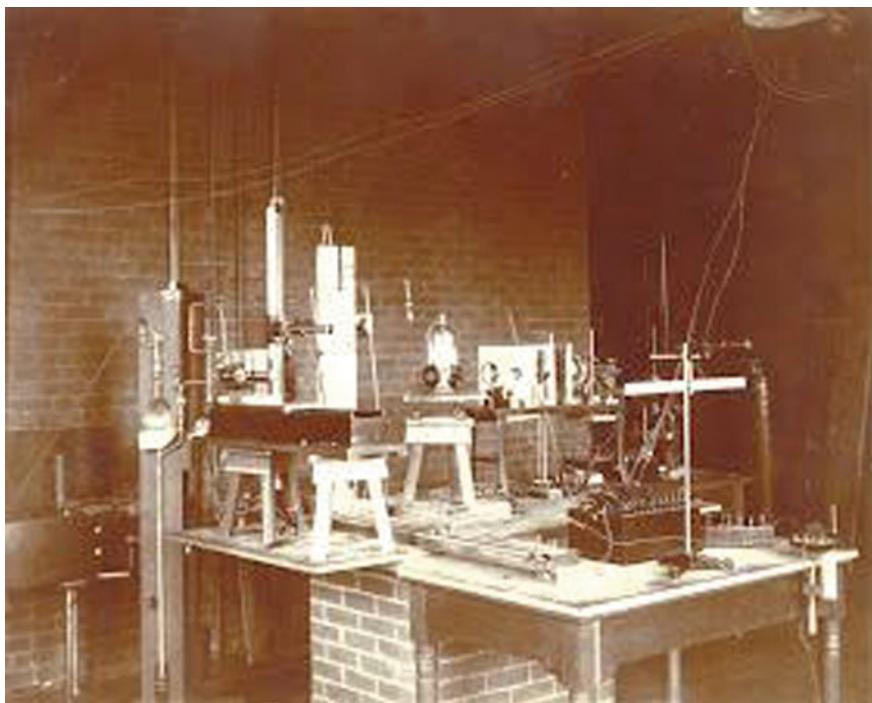


Fig. 13.1 Experimental setup for the light pressure measurements at Dartmouth College (Rauner Special Collections Library)

is attached for posterity to the Lebedev Institute of the Russian Academy of Sciences, the oldest scientific research center in Russia, founded in 1934.

Other lessons of this saga are the omnipresent danger of errors in analysis or in approximations. Because the N&H paper had a silly error of choosing the wrong logarithm, it seems a foregone conclusion that Nichols and Hull did not check each other's work. It should be standard procedure for any analyses to be double-checked by someone else, since it is so difficult to catch one's own errors. Of course, approximations must be made carefully and their implications clearly understood.

For E. F. Nichols biography see [16] (Figs. 13.1 and 13.2).



Fig. 13.2 Nichols (left) and Hull (right) outside the Physics Building at Dartmouth College, between an unnamed student. Enhanced from photograph available in Rauner Special Collections Library at Dartmouth College

References

1. P. Lebedew: Les forces de Maxwell-Bartoli dues à la pression de la lumière (P. Lebedev: The forces of Maxwell-Bartoli of light pressure), *Rapports présentés au Congrès Internationale de Physique*, **2**, 133–140 (1900), in French. See online <http://books.e-heritage.ru/book/10077976>.
2. P. Lebedew: Experiments on the Pressure Force of Light, *Ann. d. Phys.*, **4**, Is. 6, 433–458 (1901).
3. E. F. Nichols, and E. F. Hull: A preliminary communication on the pressure of heat and light radiation, *Phys. Rev.* **13**, No. 5, 307–320 (Nov. 1901).

4. P. Lebedew: Researches on the Pressure Forces of Light, *Chemical News and Journal of Physical Sciences*, **LXXXV**, p. 37, 52 and 61 (1902). Accessed 9/28/2017 at <https://books.google.com/books?id=sBILAAAAYAAJ>.
5. E. F. Nichols and G. F. Hull: The pressure due to radiation, *Proc. Am. Acad. of Arts and Sci.*, **38**, No. 20, 559–599 (1903). Available online at <http://www.jstor.org/stable/20021808>.
6. E. F. Nichols and G. F. Hull: The Pressure Due to Radiation (Second Paper), *Phys. Rev.* **17**, 26–50 (1903). This is essentially identical to the paper in ref. 5.
7. Arthur L. Day and C. E. Van Orstrand, “Black Body and the Measurement of Extreme Temperatures,” *Astrophysical J.*, vol. **19**, #1, p. 1 (1904).
8. J. H. Poynting: Radiation Pressure, *Phil. Mag.* **9**, 52, 393–406 (1905). <http://dx.doi.org/10.1080/14786440509463293>.
9. W. Gerlach and A. Golsen: A new measurement of light pressure, *Ann. d. Phys.* **73**, 1–7 (1924).
10. A. Golsen: On a new measurement of light pressure, *Ann. d. Phys.* **73**, 624–642 (1924).
11. M. Bell and S. E. Green: On radiometer action and the pressure of radiation, *Proc. Phys. Soc.* **45**, 320 (1933).
12. G. F. Hull, M. Bell and S. E. Green: Notes on the pressure of radiation, *Proc. Phys. Soc.* **46**, 301 (1934).
13. G. Burniston Brown: Resonant detection of light pressure, *Amer. J. Phys.* **34**, 272 (1966); <http://dx.doi.org/10.1119/1.1972906>.
14. M. L. Schagrin: Early observations and calculations on light pressure,” *Amer. J. Phys.* **42**, 927 (1974); <http://dx.doi.org/10.1119/1.1987899>.
15. J. Worrall: The pressure of light: the strange case of the vacillating ‘crucial experiment’, *Stud. Hist. Phil. Sci.*, **13**, No. 2, 133–171 (1982), (Pergamon Press Ltd. Gt. Britain).
16. E.L. Nichols, *Ernest Fox Nichols, A Biographical Memoir*, National Academy of Sciences, Washington, D.C. (1929).

13.2 Original Paper of E. F. Nichols and E. F. Hull “A Preliminary Communication on the Pressure of Heat and Light Radiation”

Reprinted with permission of the American Physical Society from *Phys. Rev.* **13**, No. 5, 307–320 (Nov. 1901).

A PRELIMINARY COMMUNICATION ON THE
PRESSURE OF HEAT AND LIGHT
RADIATION.

BY E. F. NICHOLS AND G. F. HULL.

M AXWELL,¹ dealing mathematically with the stresses in an electro-magnetic field, reached the conclusion that "in a medium in which waves are propagated there is a pressure normal to the waves and numerically equal to the energy in unit volume." Bartoli,² in 1876, announced that the existence of such a stress was essential to the validity of the second law of thermodynamics. He sought for such a pressure experimentally but failed to get conclusive results. The problem has been discussed theoretically, on the basis of thermodynamics, by Boltzmann³ and by Galitzine.⁴ Lebedew⁵ applies such a pressure, opposed to the sun's gravitation, to account for the solar repulsion of comets' tails; and more recently Arrhenius⁶ attempts to explain the aurora borealis on similar grounds.

Every approach to the experimental solution of the problem has hitherto been balked by the disturbing action of gases which it is impossible to remove entirely from the space surrounding the body upon which the radiation falls. The forces of attraction or repulsion, due to the action of gas molecules, are functions of the temperature difference between the body and its surroundings, caused by the absorption by the body of a portion of the rays which fall

¹ J. C. Maxwell, *Elec. & Mag.*, 1st Ed., Oxford, 1873, Vol. 2, p. 391; also 2d Ed., Oxford, 1876, Vol. 2, p. 401.

² S. Bartoli, *Sopra i movimenti prodotti della luce e dal calore*: Florence, Le Monnier, 1876; also Exner's *Repert d. Phys.*, 21, p. 198. 1885.

³ L. Boltzmann, *Wied. Ann.*, 22, p. 31; also p. 291; also *Wied. Ann.*, 31, p. 139.

⁴ B. Galitzine, *Wied. Ann.*, 47, p. 479, 1892; also *Phil. Mag.*, 85, p. 113, 1893.

⁵ P. Lebedew, *Wied. Ann.*, 45, p. 292, 1892.

⁶ S. Arrhenius, *Konigl. Vetenskaps-Akademiens, Förhandlingar*, 1900; also *Phys. Zeitschrift*, Nov. 10 and 17, 1900.

upon it, and of the pressure of the gas surrounding the illuminated body. In the particular form of apparatus used in the present study the latter function appears very complicated, and certain peculiarities of the gas action remain inexplicable upon the basis of any simple group of assumptions which the writers have so far been able to make.

Since we can neither do away entirely with the gas nor calculate its effect under varying conditions, the only hopeful approach which remains is to devise apparatus and methods of observation which will reduce the errors due to gas action to a minimum. There are four ways in which the partial elimination of the gas action may be accomplished experimentally :

1. The surfaces which receive the radiation, the pressure of which is to be measured, should be as perfect reflectors as possible. This will reduce the gas action by making the rise of temperature due to absorption small while the radiation pressure will be increased ; the theory requiring that a beam, totally reflected, exert twice the pressure of an equal beam completely absorbed.
2. By studying the action of a beam of constant intensity upon the same surface surrounded by air at different pressures, certain pressures may be found where the gas action is smaller than at others.
3. The apparatus—some sort of torsion balance—should carry two surfaces symmetrically placed with reference to the rotation axis, and the surfaces on the two arms should be as nearly equal as possible in every respect save one—the forces due to radiation and gas action should have the same sign on one side and opposite signs on the other. In this way a mean of the resultant forces on the two sides should be, in part at least, free from gas action.
4. Radiation pressure, from its nature, must reach its maximum value instantly, while observation has shown that the gas action begins at zero and increases with length of exposure, rising rapidly at first, then more slowly to its maximum effect, which, in most of the cases observed, was not reached until the exposure had lasted from $2\frac{1}{2}$ to 3 minutes. For large gas pressures an even longer exposure was necessary to reach stationary conditions. The gas action may be thus still further reduced by a ballistic or semi-ballistic method of measurement.

No. 5.]

PRESSURE OF RADIATION.

309

THE TORSION BALANCE.

The form of suspension of the torsion balance, used to measure radiation pressure in the present study, is seen in Fig. 1. The ro-

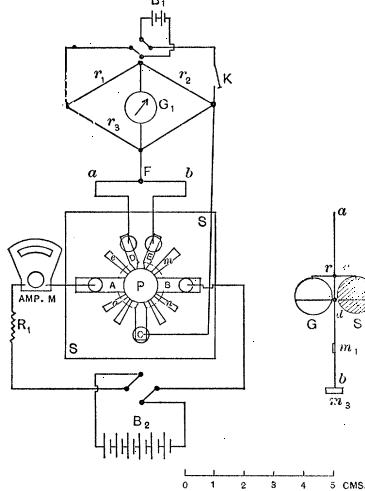


Fig. 4.

Fig. 1.

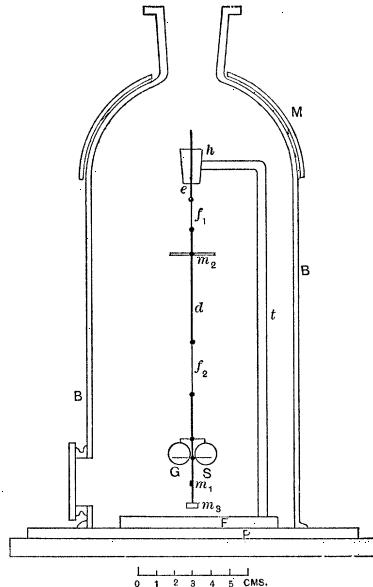


Fig. 2.

tation axis ab was a fine rod of drawn glass. At r , a drawn glass cross-arm c , bent down at either end into the form of a small hook,

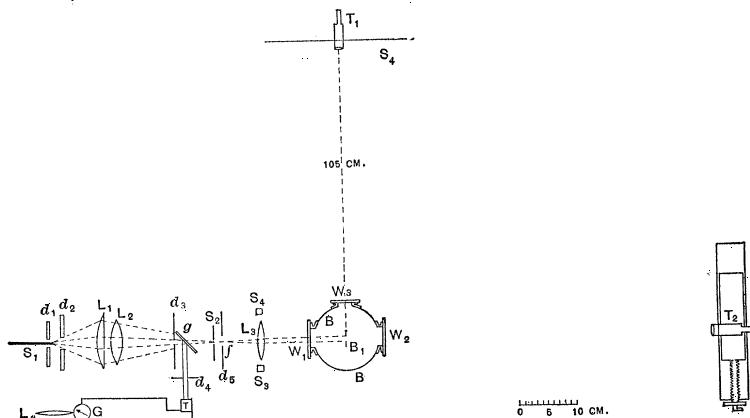


Fig. 3.

was attached. The surfaces G and S , which received the light beam, were circular microscope cover-glasses, 12.8 mm. in diameter and 0.17 mm. thick, weighing approximately 51 mgs. each. Through each glass a hole 0.5 mm. or less in diameter was drilled near the edge, by means of which the glasses could be hung on the hooks on the cross-arm c . Two other drawn-glass cross-arms were attached to the rotation axis at d , one on each side. The cover-glasses slipped easily between these, and were thus held securely in one plane. Further down on ab , a small silvered plane mirror m_1 was made fast at right angles to the plane of G and S . This mirror was polished bright on the silver side so that the scale at S_4 (Fig. 3), could be read in either face. A small brass weight m_3 (Fig. 1), of 452 mgs. mass and of known dimensions, was attached at the lower end of ab . The cover glasses which served as vanes were silvered and brilliantly polished on the silvered sides, and so hung on the small hooks on c that the glass face of one, and the silver face of the other, were presented to the light. A quartz fiber f_2 (Fig. 2), 3 cms. long, was made fast to the upper end of ab , and to the lower end of a fine glass rod d which carried a horizontal magnet m_2 . The rod d was in turn suspended by a short fiber to a steel pin e which could be raised or lowered in the bearing h . The whole was carried by a bent glass tube t , firmly fastened to a solid brass foot F , resting on a plane, ground-glass plate P , cemented to a brass platform mounted on three levelling screws not shown in Fig. 2. A bell-jar B , 25 cms. high and 11 cms. in diameter covered the balance. The flange of the bell-jar was ground to fit the plate P . A ground-in hollow glass stopper fitted the neck of the bell-jar, which could thus be put in connection with a system of glass tubes leading to a Geissler mercury pump, a MacLeod pressure gauge, and a vertical glass tube dipping into a mercury cup and serving as a rough manometer for measuring the larger gas pressures employed during the observations. The low pressures were measured on the MacLeod gauge in the usual way. A semicircular magnet M , fitted to the vertical curvature of the bell-jar, was used to direct the suspended magnet m_2 , and thus control the zero position of the torsion balance. By turning M through 180° , the opposite faces of the vanes G and S could be presented to the light.

THE ARRANGEMENT OF APPARATUS.

A horizontal section of the apparatus through the axis of the light beam, is shown in Fig. 3, which like the other figures is, in its essential features, drawn to scale. The white-hot end of the horizontal carbon S , of a 90° arc lamp, fed by alternating current, served as a source. The arc played against the end of the horizontal carbon from the vertical carbon which was screened from the lenses L_1 and L_2 by an asbestos diaphragm d_2 . The lens L_4 projected an enlarged image of the arc and carbons on a neighboring wall, so that the position of the carbons and the condition of the arc could be seen at all times by both observers. The cone of rays passing through the small diaphragm d_2 fell upon the glass condensing lenses L_1 , L_2 which formed an image of the carbon at f . At d_3 a diaphragm, 11.25 mm. in diameter, was interposed, which permitted only the central portion of the cone of rays to pass. Just beyond d_3 , a plane plate of glass g was placed in the path of the beam and reflected part of the beam through the diaphragm d_4 , to a thermopile T , connected with a galvanometer G . The beam which traversed the plate g passed on to a shutter at S_2 . The shutter was worked by a magnetic escapement, operated by the seconds contact of a standard clock. The observer at T_1 might choose the second for opening or closing the shutter, but the shutter's motion always took place at the time of the seconds contact in the clock. Any exposure was thus of some whole number of seconds duration. The opening in the shutter was such as to let through, at the time of exposure, all of the direct beam which passed through d_3 , but shut out the stray light. It was made to correspond in every way to the diaphragm d_4 , so that for all of the reflected beam which passed d_4 , the corresponding transmitted portion passed S_2 and no more. The glass lens L_3 focused a sharp image of the aperture d_3 in the plane of the vanes of the torsion balance B_1 under the bell-jar B . The bell-jar was provided with three plate glass windows W_1 , W_2 , W_3 . The first two gave a circular opening 42 mm. in diameter, and the third was used in connection with the telescope T_1 , by which an image of the scale S_4 was seen in the small mirror m_1 (Figs. 1, 2). The lens L_3 was arranged to move horizontally between the stops

S_3 and S_4 . The stops were so adjusted that when the lens was against S_3 the sharp image of the aperture d_3 fell centrally upon one vane; and when against S_4 the image fell centrally upon the other. This adjustment, which was a very important one, was made by the aid of a telescope T_2 , mounted on the carriage of a dividing engine. This was used to observe and measure the positions of the vanes and the rotation axis, as well as the positions of the images of d_3 , when the lens L_3 was against the stops S_3 and S_4 . For the latter measurements, the vanes could be moved out of the way, by turning the suspension through 90° by the control magnet M (Fig. 2).

METHODS OF OBSERVATION.

The observations leading to the results given later were of three different kinds: (1) The calibration of the torsion balance; (2) the measurement of the pressure of radiation in terms of the constant of the balance, and (3) the measurement of the energy of the same beam in electrical units by the aid of a bolometer described later.

1. The determination of the constant of the torsion balance was made by removing the vanes G and S and accurately measuring the period of vibration. Its moment of inertia was easily computed from the masses and distribution of the separate parts about the axis of rotation. From the moment of inertia, period, the lever arms of G and S , and the distance of the scale S_4 , the constant of the balance for a centimeter deflection was found to be 4.65×10^{-5} dynes.

2. In the measurements of radiation pressure, it was easier to refer the intensity of the beam at each exposure to some arbitrary standard which could be kept constant than to try to hold the lamp as steady as would otherwise have been necessary. For this purpose, the thermopile T (Fig. 3) was introduced, and simultaneous observations were made of the relative intensity of the reflected beam by the deflection of the galvanometer G , and the pressure due to the transmitted beam by the deflection of the torsion balance. The actual deflection of the balance was then reduced to a deflection corresponding to a galvanometer deflection of 100 scale divisions. The galvanometer sensitiveness was carefully tested at

No. 5.]

PRESSURE OF RADIATION.

313

intervals during the work, and was found as nearly constant as the character of the observations required. All observations of pressure were thus reduced to the pressure due to a beam of fixed intensity.

At each gas pressure in the bell-jar, at which radiation pressure measurements were taken, two sets of observations were made. In one of these sets, static conditions were observed, and in the other, the deflections of the balance due to short exposures were measured. In the static observations, each vane of the balance was exposed in turn to the beam from the arc lamp, the exposures lasting in each case until the turning points of the swings showed that static conditions had been reached. The combined pressure, due to radiation and gas action would thus be equal to the product of the angle of deflection and the constant of the balance. The torsion system was then turned through 180° by rotating the outside magnet, and similar observations made on the reverse side of the vanes, after which the system was turned back to its first position, and the earlier set of observations repeated, and so on.

The considerations underlying the above method of procedure, are as follows: The beam from the lamp, before reaching the balance, passed through three thick glass lenses and two glass plates. All wave-lengths destructively absorbed by the glass, were thus sifted out of the beam by the time it reached the glass balance vanes. The silver coatings on the vanes absorbed therefore more than the glass. The radiation pressure was always away from the source irrespective of the way the vanes were turned, while the gas action would be exerted mainly on the silvered side of the vanes. As the face of the vane on one side of the balance was silver, and on the other side glass backed by silver, the two forces should act in the same direction on one vane, and oppositely on the other. If each silver coating reflected equally on its two faces, the average of the deflections on the two arms of the balance, should be free in part, from gas action.

This partial elimination of gas action would be more complete but for two reasons: (1) The reflection at the exposed glass face of one vane diminishes the energy incident upon the silver coating behind it and the resultant absorption in the same propor-

tion. (2) The glass faces of the two vanes, when exposed, will not, in all probability, be equally heated. In the case of the vane which receives the light on its silvered face, the glass face can be warmed only by the conduction of heat from the silver coating through the glass, while in the vane receiving light on its glass face, some heat from absorption of the beam in passing through the glass will doubtless be added to this heat due to conduction. Although the vanes, when freshly silvered and polished, were so brilliant on the silvered sides that it was often difficult to tell which face was silver and which glass, the coatings were found in use to deteriorate rapidly, and the silver faces whitened earlier than the glass. After several hours of observing, vanes freshly silvered at the beginning had to be removed and resilvered. A higher rate of deterioration was noticeable when working in low gas pressures than in high.

By the reversal of suspension and averaging the results gained with the suspension direct and reversed, nearly all errors due to lack of symmetry in the balance, or in the position of the light images with reference to the rotation axis; or errors due to lack of uniformity in the distribution of radiant intensity in different parts of the image, as well as errors due to certain differences between the mirrors themselves, could be eliminated.

Static observations were made at eight different gas pressures between 0.06 mm. of mercury and 96 mm. Within this range, as was afterwards shown by the series of semiballistic in connection with the static observations, the gas action varied in magnitude between one-tenth, and 6 or 7 times, the radiation pressure. The gas action was, however, at certain pressures so complicated that no rational way of treating the static measurements at the different gas pressures succeeded in establishing the existence of a constant force in the direction of propagation of the light beam. It was plain, therefore, that further elimination of the gas action must be sought in exposures so short that the gas action would not have time to reach more than a small fraction of its stationary value. This led to the method of semiballistic observations.

If the vibrations of any suspension be undamped, a constant deflecting force, acting for one-fourth of the period of oscillation, will

No. 5.]

PRESSURE OF RADIATION.

315

cause a throw equal to $\sqrt{2}$ times the permanent deflection for the same force. A correction for damping when considerable may also be introduced. The period of the balance as described was 23.1 secs. This was increased to 24 secs. by adding masses to the system. Deflections due to six seconds' exposure were then observed in the same manner, with reversals and all, as has been already described under the static observations. Because of the shorter time involved in making individual observations, a considerably greater number of ballistic measurements were made at each gas pressure. In all the observations yet made, not a single ballistic deflection contrary to the direction of radiation pressure was obtained, and in the static observations, the first throw on exposing was invariably in the radiation pressure direction, irrespective of the direction of the final deflection.

The average ballistic deflection obtained, reduced to the equivalent permanent deflection for each of eight different gas pressures, is given in the accompanying table. The column p gives the gas pressure in the bell-jar in millimeters of mercury, and column d the reduced permanent deflection in millimeters for the radiation pressure.

p	d	p	d
96.3	19.7	33.4	21.1
67.7	21.0	1.2	20.9
37.9	21.6	0.13	26.8
36.5	22.1	0.06	23.2
Ave.			22.5

For gas pressures between 45 and 30 mm. the gas action is small and increases toward both higher and lower pressures in the range studied. Taking the product of the average deflection in cms. by the constant of the torsion balance, the radiation pressure was $2.25 \times 4.65 \times 10^{-5} = 1.05 \times 10^{-4}$ dynes.

THE BOLOMETER.

To test the above value of radiation pressure with the theory, it was necessary to measure the energy of the radiation causing the

pressure. This was attempted with the aid of a bolometer of special construction.

On a sheet of platinum 0.001 mm. thick, rolled in silver (by the firm Sy & Wagner, Berlin), a circle P (Fig. 4), 11.3 mm. in diameter, was drawn. The sheet was cut from the edges inward to the circumference of the circle, in such a way as to leave five principal strips A , B , C , D , E , connected to the circle in the manner shown. Other narrower strips, as e , m , n , o , etc., were left to give the disk additional support. The disk, by means of the connecting arms, was mounted with asphalt varnish centrally over a hole 14 mm. in diameter, bored through a slab S cut from a child's school slate. Portions of the silver not to be removed by the acid were carefully covered by the asphalt varnish. Thus on the strips A and B , the silver was protected to the very edge of the circle, while on all the other arms, the silver was left exposed back to the edge of the boring in the slate. The whole system was then plunged into warm nitric acid, and the silver eaten away from all unprotected surfaces, leaving only the thin platinum sheet, which was blackened by electric deposition of platinum, by the method used in the manufacture of the Reichsanstalt bolometers. At A , B , C , D , E , holes were bored extending through the slate, and copper washers were soldered to the silver strips, and binding posts were attached.

The torsion balance was removed from under the bell-jar, the bolometer put in the place of one of the vanes, and the bell-jar replaced. Connections to the bolometer were made as schematically shown in Fig. 4. The disk P was the exact size of the light image thrown on the vanes in the pressure measurements, and the intention was to heat the disk by allowing the image to fall on it, and then with the light turned off, to heat it to the *same temperature* by sending a current through it from A to B . If r be the resistance from A to B in ohms, when exposed to the lamp, and i be the current in ampères which gives the same temperature in P as that given by the absorbed radiation, then $i^2r \times 10^7$ will be the activity of the beam in erg-seconds. The temperature of the disk, whether exposed to the radiation or heated by the current, was shown by the resistance of the disk in the direction C to D and E which was

No. 5.]

PRESSURE OF RADIATION.

317

made one arm of a Wheatstone Bridge. The relation of the heating current to the bridge was adjusted as follows: With the key K open, so that no current flowed through the bridge, the heating current from six storage cells B_2 was turned on, and the sliding contact at F so set that the bridge galvanometer zero was not changed by reversing the heating current. The equipotential point to c was found very near the middle of the wire ab , which showed the current distribution of P to be symmetrical with respect to a diameter at right angles to AB . The key K was then closed making the bridge current, and the bridge was balanced. The bolometer was next exposed to the radiation, and simultaneous observations were made on the two galvanometers G (Fig. 3) and G_1 (Fig. 4) and the deflection of the latter, was reduced to that corresponding to a deflection of 100 divisions of galvanometer G , as was done in the radiation pressure observations. During the pressure, as well as the bolometer measurements, the electric lamp was regulated by the observer at the galvanometer G , and the lamp held reasonably close to this standard value. The reductions were, therefore, of the simplest kind. After shutting off the light, the heating current was turned on and regulated by means of the variable resistance R_1 (Fig. 4), so that nearly the same throw was obtained from the galvanometer G , as when the bolometer was exposed to the lamp. All deflections of the galvanometer G_1 were taken with the bridge current both direct and reversed, to eliminate any local disturbances in the bridge. With the heating current on, galvanometer deflections were read with the heating current both direct and reversed. The heating current was read in ampères from a Siemens & Halske direct-reading precision milliampèremeter. From repeated observations the current which gave the same heating effect as the light beam, was $i = 0.865$ amp. The resistance between the binding posts A and B , was measured with the lamp on, and gave $r = 0.278$ ohm. The intensity of the beam in erg-seconds was thus $= 0.278 \times 0.75 \times 10^7$.

Considering the beam as a cylinder containing energy which is transmitted at the velocity v of radiation, the energy contained in a length of 3×10^{10} cms. would, according to the bolometer results be $i^2 r \times 10^7$, and the consequent pressure in dynes, against the totally

absorbing end of the cylinder must, according to Maxwell, equal the energy per unit length, or

$$P = \frac{i^2 r \times 10^7}{v}.$$

If the incident beam falls upon a totally reflecting surface the energy per unit volume is doubled and the deduction from Maxwell is that the pressure is equal to $\frac{2i^2 r \times 10^7}{v}$.

According to Bartoli and Boltzmann, the pressure due to a beam totally reflected must on thermodynamic grounds also be $\frac{2i^2 r \times 10^7}{v}$.

If the beam is but partially reflected, however, the pressure would be

$$P = \frac{(1 + r_1)i^2 r \times 10^7}{v},$$

where r_1 equals the percentage of the incident energy reflected. Considering the character of the light from the arc lamp, and the heavy infra-red absorption of the glass masses traversed, the reflecting power of the bright silver coatings was estimated to be nearly that for the wave-lengths of the *D* lines, or about 92 per cent.¹ Putting $1 + r = 1.92$ we have from the bolometer results

$$\frac{1.92 \times 0.278 \times 0.75 \times 10^7}{3 \times 10^{10}} = 1.34 \times 10^{-4} \text{ dynes}$$

as the pressure which the torsion balance should have shown, instead of the observed value 1.05×10^{-4} . Some of the mirrors used in the pressure observations were already well below this maximum reflecting power, so that the value 1.34×10^{-4} should doubtless be still further reduced, but no quantitative value can be assigned to such a correction.

Taking the results as they stand, the measured radiation pressure is to the radiation pressure which the theory applied to the bolometer measurements would require as $1.05 : 1.34$ or as $78 : 100$.

¹ E. F. Nichols, Phys. Rev., 4, p. 303 also F. Paschen, Ann. Phys., 4, p. 304, 1901.

Some sources of uncertainty in the pressure measurements have already been mentioned, but certain sources of error in the bolometer measurements in which the nature of the effect upon the result is known, should be more clearly indicated.

1. Errors in the bolometer which would tend to produce a greater divergence between the two results obtained are :

(a) The reflection of radiation by the platinum-black coating (this error is small) and (b) the heat conductivity of the black coating. In the case of the lamp, the transformation of radiant energy into heat takes place near the outer surface of the platinum-black covering, and the electrical resistance of P is affected only after the heat has passed by conduction through this coating to the sheet of platinum underneath. In the case of heating by the current, on the other hand, the heat is generated in the conducting sheet directly, and is dissipated at the outer surface after conduction through the black layer. This error is probably small also.¹

2. Errors which would tend to bring the two results into closer agreement are two in number. (a) The measured resistance between the binding posts A and B must be greater than that of the disk P by the resistance of the silver strips leading from the binding posts on both sides to P , and by any contact resistances involved. The magnitude of this error is unknown, but it may be large. (b) The distribution of current lines of flow for both heating current and bridge current would be such as to cause an error in the result of the same sign as the resistance error. The maximum heating would occur where the silver strips A and B join the disk, for here the stream lines will be most congested; on the other hand, the stream lines of the bridge current in these regions will be most sparse because nearly all the lines traversing the disk on the two sides will be drawn off into the edges of these same silver strips by the higher conductivity of the latter. This error may also be considerable. In a new bolometer now under construction the errors due to false resistance and unfavorable distribution of flow lines will be greatly reduced. Another series of radiation pressure measure-

¹ See Kurlbaum: Temperaturdifferenzen zwischen der Oberfläche und dem Innern eines strahlendes Körpers. Ann. Phys., 2, p. 554.

320

E. F. NICHOLS AND G. F. HULL.

[VOL. XIII.]

ments, made under more favorable conditions than those here reported, is also in progress.

The writers believe that the observations already in hand are sufficient to prove experimentally the existence of a pressure, not due to gas molecules, of the nature and order of magnitude of radiation pressure, but toward a close quantitative measurement of this pressure much remains still to be done.

THE WILDER LABORATORY, DARTMOUTH COLLEGE,
HANOVER, N. H., August, 1901.

Chapter 14

The First Nonlinear Optical Experiment of 1926, Measuring Sensitivity Threshold of the Human Eye to Feeble Light (1933) and Statistical Structure of Feeble-Light Interference by the Human Eye (Sergei Ivanovich Vavilov)



Svetlana G. Lukishova

Abstract This chapter is dedicated to the scientific contributions to photonics of the legendary figure in the history of the Soviet/Russian science, Sergei Ivanovich Vavilov (1891–1951). Among his key contributions to photonics are: (1) first experimental observation of nonlinear optical effect (1926), coining the terminology “nonlinear medium” in optics, “nonlinear interferometry” (1930), and “nonlinear optics” (1944); (2) observation of quantum fluctuations of light by the human eye excluding physiological fluctuations (measuring a sensitivity threshold of the human eye to the feeble light) (1933–1942). Statistical structure of interference pattern was studied by Vavilov in light fleshes at mean photon numbers close to the human eye sensitivity threshold. Independent fluctuations in coherent split beams were also observed; (3) decisive contributions to the discovery of Vavilov-Čerenkov radiation known in the West as Čerenkov radiation; (4) transferring the field of luminescence to the scientific platform, starting from its definition, introducing the concept of fluorescence quantum yield, polarized fluorescence and Vavilov’s luminescent law. The chapter consists of five sections. In the first Sect. (14.1) by Svetlana G. Lukishova a detailed analysis of Vavilov’s contribution to science of light is provided including his organizing and administrative work at the State level as well as some biographical material. On editor selection, Sects. 14.2–14.4 contain translations into English of some of his publications and comments. Section 14.5 contains original publication of Vavilov et al. in English.

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road,
Wilmot Building, Rochester, NY 14627, USA

e-mail: sluk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

14.1 Sergei Vavilov's Nonlinear and Quantum Optical Experiments in Pre-laser Times, Vavilov-Čerenkov Radiation, and His Contributions to the Luminescence Science

14.1.1 Introduction

This part of the Chap. 14 is dedicated to the scientific contributions to photonics of the legendary figure in the history of the Soviet/Russian science, Sergei Ivanovich Vavilov (Wawilow) [1–22] (Fig. 14.1). On 24 March 2016 was the 125 anniversary of his birthday. Among his key contributions to photonics are: (1) first experimental observation of nonlinear optical effect (1926), coining the terminology “nonlinear medium” in optics, “nonlinear interferometry” (1930), and “nonlinear optics” (1944); (2) observation of quantum fluctuations of light by the human eye excluding physiological fluctuations (measuring a sensitivity threshold of the human eye to the feeble light) (1933–1942). Statistical structure of interference pattern was studied by Vavilov in light fleshes at photon numbers close to the human eye sensitivity threshold. Independent fluctuations in coherent beams were also observed (1933–1934); (3) decisive contributions to the discovery of Vavilov-Čerenkov radiation known in the West as Čerenkov radiation; (4) transferring the field of luminescence to the scientific platform, starting from its definition, introducing the concept of fluorescence quantum yield, polarized fluorescence and Vavilov's luminescent law.

In addition to his personal scientific contributions, it is valid to say that S. I. Vavilov was the soul and organization engine of Soviet optical science throughout pre-Word-War-II, the War, and post-War years [3]. He is counted as the great builder of science in the USSR [3]. He was the founder of the P. N. Lebedev Physical Institute in Moscow [4] (see Fig. 14.2) and at the same time the director for research of the Optical State Institute (GOI) in Leningrad (St.-Petersburg today) [5] sharing his time between these two cities. His life reminds Shakespeare tragic players: his fate was to be the President of the Soviet Academy of Sciences in one of the harshest periods of Stalin's time (1945–1951): his brother Nikolai I. Vavilov, the world known plant explorer, seed collector, and genetic scientist who formulated the law of homologous series variability was imprisoned in 1940 and died in prison from starvation in 1943.

In spite of his intensive organizing and administrative work at the State level, he left us many scientific papers and books written him personally. He liked to say to his students and co-workers that “*ars longa, vita brevis* (art is long, life is brief)” [3]. He was very hard working person and very well organized. The Soviet Academy of Sciences published four volumes of his publications [6]. These volumes comprise his own original papers of 1914–1936 (vol. I, 50 papers, 398 p.), his original papers of 1937–1951 and the book of 1950 “Microstructure of Light” (vol. II, 34 papers and book, 547 p.), his papers on the philosophy and history of natural science (vol. III, 56 papers, 870 p.), including biographies, theories and description

Fig. 14.1 S. I. Vavilov
(1936)

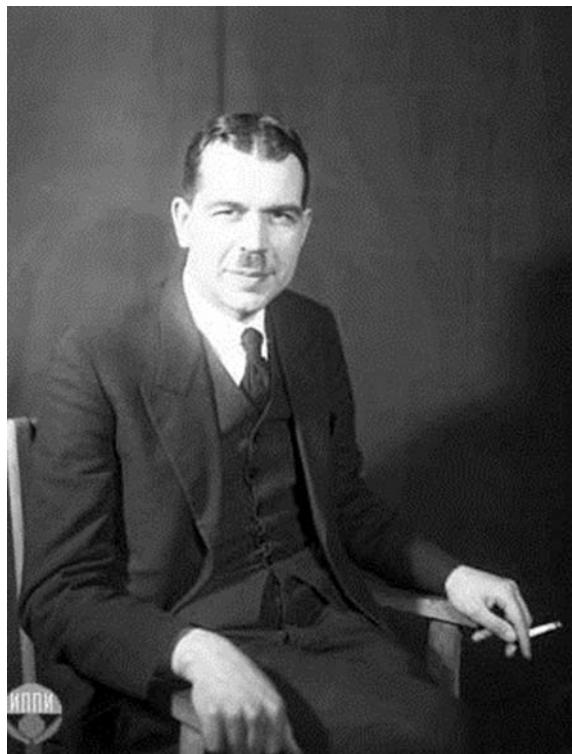


Fig. 14.2 Vavilov in his Lebedev Physical Institute office with P. N. Lebedev's devices on measurement of the light pressure (1947)



of the work of great scientists and philosophers (Lucretius, Newton, Galileo, Huygens, Grimaldi, Faraday, Michelson, Lomonosov, Euler, Lebedev, Lazarev, Krylov, etc.) as well as the history of Russian science through 220 years, Soviet science during its 30 years, French revolution science and technology, etc.

Volume IY (470 p.) contains his three books “Experimental Foundations of the Theory of Relativity” (101 p.), “About “Warm” and “Cold” Light” (48 p.) and “The Human Eye and the Sun” (96 p.). Volume IY also contains his 12 review and popular papers as well as his 4 articles for the encyclopedias. Vavilov was also an Editor-in-chief of the second edition of the Great Soviet Encyclopedia. He read every article of it and edited articles himself if he was not satisfied by them [11]. His former student, 1958 Nobel prize winner I. M. Frank compares his personality with very wide interests not only to science, but the fine art, literature, and foreign languages (Vavilov knew Latin, German, English, French, Italian, Polish), like the Italian Renaissance personality [3].

At the same time another 1958 Nobel prize winner I. E. Tamm writes about S. I. Vavilov: “He was the God-gifted physicist and understood the music of all physics”. And many people knowing him emphasize his human personal qualities. Russian literature contains many sincere reminiscences about him, e.g., [3–5, 7–19], see, e.g., translated into English the chapter of the book [7] of E. L. Feinberg who knew S. I. Vavilov well working at the Lebedev Physical Institute and who thought a lot about Vavilov’s personality. One of Vavilov’s colleagues from the State Optical Institute [15] wrote about him, that “On the shoulders of such people as Sergej Vavilov, “the Earth can be reliably held”. Such people leave a very deep trace after them, on their life examples young generations are educated, and the memory about them piously hold in their hearts all who had the good luck to work with them”.

Some biographical events and some of his publications are listed on the website of the S. I. Vavilov Institute for the History of Science and Technology of the Russian Academy of Sciences [13]. Lebedev Physical Institute’s website contains the book [12] about him. The Russian Academy of Sciences prepared a CD devoted to his 120 year anniversary [20]. Unfortunately, most of them were not translated from Russian, but we can recommend [7–9, 14, 16] in English with some reminiscences about him. References [1–2, 4, 8, 9, 14, 16, 18–19, 21–24] in English contain information about his scientific contributions.

The present Sect. 14.1 is devoted to S. I. Vavilov’s scientific contributions to photonics (in the wider sense of this direction of science). It is organized as follows. Section 14.1.2 describes first nonlinear optical experiment known to the author which was carried out in 1925 (published in 1926) by Vavilov and Levshin [25]. This section also contains an outline of Vavilov’s 1951 book “Microstructure of light” [26] where he writes about nonlinear optics. Section 14.1.3 is devoted to Vavilov’s and Brumberg’s method of measurement of sensitivity of a human eye to a small number of photons (observation of quantum fluctuations of light by a human eye) and studying statistical properties of coherent light beams and interference pattern by a human eye near threshold sensitivity [26–35]. Section 14.1.4 describes the history of Vavilov-Čerenkov radiation and Vavilov’s decisive role in this discovery [36, 37]. Vavilov’s contributions to the luminescence science [18, 19] and his creation of schools of luminescence [18, 19] at the Lebedev Institute (Moscow) and the State Optical Institute (GOI) (Leningrad), where later first Soviet lasers were made, are discussed in Sect. 14.1.5. Section 14.1.6 describes Vavilov’s

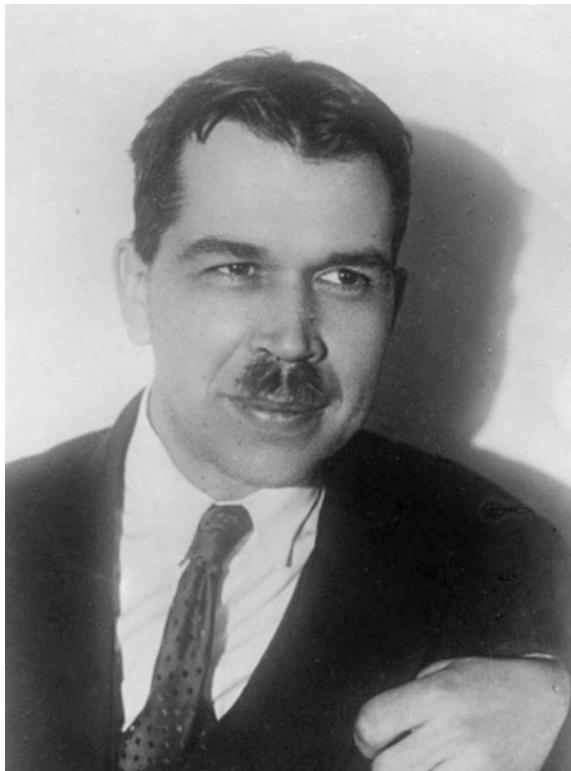
works on the history of Science and its popularization. Vavilov's organizing and administrative activities as the founder and director of Lebedev Physical Institute in Moscow and the director for research of the State Optical Institute as well as President of the Soviet Academy of Sciences are briefly described in Sect. 14.1.7. Section 14.1.8 contains some biographical information about S. I. Vavilov and his family.

14.1.2 Vavilov's Nonlinear Optical Experiment (1926), Coining “Nonlinear Optics” Terminology and His Book “Microstructure of Light”

Thirty five years before the first laser, on December 1925, Vavilov and Levshin submitted paper on observation of the first ever nonlinear optical effect using a pulsed condensed spark light source [25], see also Vavilov's book [26] and Sect. 14.2 with translation of excerpts from [25]. With increasing incident energy, Vavilov and Levshin found diminishing the absorption of light by a glass doped with uranium. Saturation of the absorption in a medium with a long-lived excited state ($\sim 5 \cdot 10^{-4}$ s) was the reason for this effect. This experiment violated Bouguer's (Beer–Lambert–Bouguer) law which states on independence of light absorption coefficient on power or intensity. These results of [25] of Vavilov and Levshin on fluorescence and phosphorescence measurements were presented by Peter Pringsheim on 12 May 1926 [see Vavilov's letter to Levshin in [11], p. 356] at the Berlin Colloquium during Vavilov's visit of a Pringsheim's laboratory in Berlin. Pringsheim started from an overview of the uranium glass experiments [although it was a subject of a last paragraph (§ 4) of this paper] and its content was accepted very well by participants [among listeners were Einstein (who asked where this work was done), Nernst, Planck (who visited Vavilov's laboratory in Moscow in 1925), Laue, Bethe, Bose, Ladenburg)]. On May 17 1926 Vavilov presented this work in Göttingen.

This experiment was a continuation of Vavilov's work started in 1918. In his earlier experiments using Rhodamine B dye doped into gelatin films with a short-lived excited state ($\sim 10^{-9}$ s), Vavilov verified Bouguer's law over a large range of incident light intensities, $2.5 \cdot 10^{-12} - 2 \cdot 10^8 \frac{\text{erg}}{\text{cm}^2 \cdot \text{s}}$ [26, p. 68], and in 1920 he published the results with 1–2% accuracy that a violation of Bouguer's law was not observed [38, 39] both at high and at low light intensities. At low light level Vavilov did not observe any quantum fluctuations in [38–39], that was his purpose, and made a conclusion that these results contradict quantum theory of M. Planck. It should be mentioned, that experiments of 1920 [38–39] were made by Vavilov at a time when quantum theory was in the beginning of its development and even Planck still made some classical assumptions in his theory of a black-body radiation in his book of 1913, see [40]. Figure 14.3 shows Vavilov approximately in 1920 year [11].

Fig. 14.3 S. I. Vavilov
(approximately in 1920)



In his book [26], pp. 68–69, Vavilov reinterprets his experiments of 1920 at very low light intensities [38–39], lower than $1 \frac{\text{photon}}{\text{cm}^2\text{s}}$, with a human eye as a detector. He explained these first experiments on quantum fluctuations of light at this time by a physiological property of the human eye (keeping image during approximately 0.1 s) and averaging actions of many photons in a continuous in time light beam by the human eye and the brain. In later experiments of 1933–1942 on quantum fluctuations of light, to avoid this problem, Vavilov used a short-time-duration light irradiation, small size of image on a retina, and the eye fixation technique.

After the experiments of 1918–1920 years, in his 1926 year paper [25], Vavilov considered again the question of the linearity of absorption suggesting about possible deviation from it at high light intensities connected with the duration of the excited state lifetime. In this case he decided to use an optical medium with a relatively long excited state lifetime ($\sim\text{ms}$) with a possibility of accumulation of an appreciable fraction of the centers in the excited state. In this case Vavilov and Levshin showed a decrease in absorption coefficient of uranium glass at high incident energy by 1.5% with a mean error of $\pm 0.3\%$. In these experiments, light from a condensed spark source with pulse duration near 10^{-6} s, 900 erg energy, was concentrated into uranium glass (0.16 cm^2 area), so that the incident intensity in the

center reached the value $\sim \text{kW/cm}^2$. One half of the spark image passed through uranium glass, the other half passed through the air. After that this image was projected onto two slits of the spectrophotometer. In the book [26, p. 72] Vavilov mentioned that a König-Martens spectrophotometer [41, 42] was used in these experiments. In this spectrophotometer the observers compared the brightness of two channels. The measurements were made at the wavelength of 454 nm. Attenuating the light was accomplished by yellow glass filter with 3 times attenuation at this wavelength. To measure the dependence of absorption on incident energy, light was sent through either (i) the transmission filter and then uranium glass or (ii) uranium glass and then the same filter. For each position of the filter, 50 measurements were carried out, and an observer did not know in which position the glass filter was placed [25, 26]. In addition, similar experiments were carried out with an aqueous solution of fluorescein dye with fluorescence lifetime $\sim 5 \cdot 10^{-9}$ s. *Under the same conditions with the same accuracy no positive results were observed with fluorescein dye in aqueous solution.*

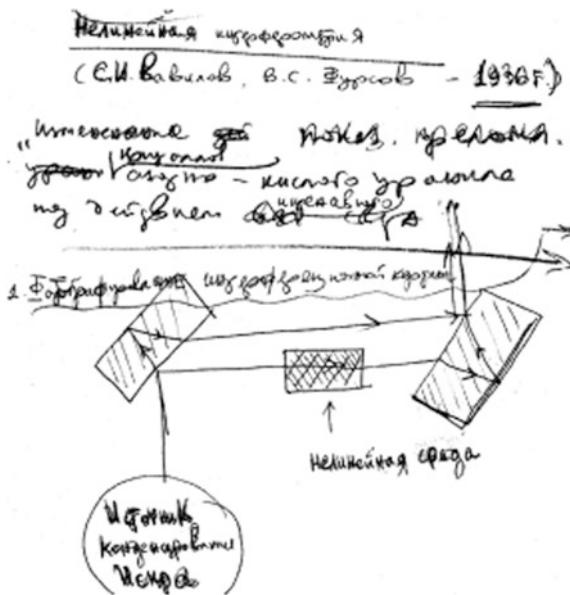
The clearly marked nonlinear phenomena in crystals with phosphorescence ("phosphors"), where the lifetimes of the excited states may be especially long, permitted K. B. Panshin to realize Vavilov's idea of building a certain type of absolute nonlinear photometer not requiring a standard comparison source (Fig. 21 of [26]).

In 1941, a similar nonlinear optical effect was reported by famous US chemist G. N. Lewis and his co-workers D. Lipkin and T. T. Magel [43]. See Chap. 15 of the current book. They reported in [43] saturation of phosphorescence and absorption with increasing the excitation intensity using the dye fluorescein, but in a boric acid glass host. Lewis et al. used much brighter light source (the 1000 W high-pressure mercury arc) with a thermostat for the sample to avoid heating effects of a powerful light. See also modern [44–45] discussing how low can be saturation threshold. In [45] saturation effect in transmission of fluorescein in a boric acid glass was observed at intensity 15 mW/cm^2 that much lower than in Vavilov and Levshin's experiment. During the Second World War US optical scientist B. O'Brian and his associates used strong saturation of the luminescence in their phosphores in the device which became known as Icaroscope to resolve problem of blinding by the sun defenders of US Navy ships against Japanese bombing attacks [44, 46]. See Chap. 16 of the current book.

For Vavilov and his students the next step after the saturation effect in absorption was the measurement of the intensity dependence of the refractive index of uranium glass. Figure 14.4 shows the schematics of this experiment. It was taken from S. A. Akhmanov's lecture on the history of nonlinear optics at Moscow State University for the XII International Conference on Coherent and Nonlinear Optics, Moscow 1985 (see also [23]). The header is translated as follows: "Nonlinear interferometry" (S. I. Vavilov, V. S. Fursov—1930 year), "Changing the refractive index of a nitrogen oxide uranyl crystal under intensive light".

The header is followed by a sketch of a nonlinear interferometer with nonlinear medium in which a spark light source is also indicated. The subtitle above the sketch of the experiments tells: "I. Photographing the interference pattern".

Fig. 14.4 Notes with suggested experiment of 1930 year on changing refractive index under intensive light irradiation (in Russian), from the lecture of S. A. Akhmanov. See explanations in the text



Vasilii S. Fursov, Vavilov's student at Moscow State University at this time and later the Dean of Physics, was asked by Vavilov to set up an interferometric experiment to detect the light intensity dependence of the refractive index with a spark source, but he did not observe any changes of refractive index [23] at intensities that were available at this time. After laser invention, in 1962, at Lebedev Institute G. A. Askaryan suggested the phenomenon of self-focusing and Hercher at the University of Rochester in 1964 and Pilipetsky and Rustamov at Moscow State University in 1965 first observed this phenomenon (see book [47]).

The terminology “*nonlinear optics*” was introduced by Vavilov. In 1944 he wrote: “We should not forget that the real optics of the matter with which we deal is nonlinear in the general cases, and its treatment requires a “nonlinear” mathematical apparatus” [18]. In his book of 1950 [26] with the title in Russian “*Mikrostruktura Sveta*” (“Microstructure of Light”) he used this terminology. See also [19, 23] with English translation of two paragraphs from this book, where Vavilov wrote:

““Nonlinearity” in an absorptive medium should be observed not only with respect to absorption. The last is connected to dispersion, therefore, the speed of light in the medium, generally speaking, should also depend on the light’s power. For the same reason, in the general case, a dependence on the light power, i.e., a violation of the superposition principle (see [48]), should be observed in other optical properties of the medium, such as birefringence, dichroism, rotation of polarization, etc.

With “*nonlinear optics*” of the medium an astrophysicist deals constantly under theoretical consideration of conditions inside stars. Because of the enormous density of light energy inside the stars at temperatures of many millions of degrees, the

absorption and the speed of light must depend on the light power very sharply. However, as mentioned above, an optical scientist also observes large violations of superposition principle and nonlinearities in a modest laboratory conditions, especially in the study of phosphorescent media.

However, physics is so accustomed to the linearity of everyday optics that until now no rigorous formal mathematical apparatus has been developed for solving real “nonlinear” optical problems...”

Vavilov was also pretty close to the problem of obtaining coherent radiation from the macroscopic ensemble of emitters in his thoughts about elementary processes of light emission. Ten years before the first laser he wrote in his book [26] in a paragraph about light beam interference (p. 78): “Is the possibility excluded of obtaining a coherent light during a sufficiently long time interval from two different particles of the matter which are located at the distances of the order of several particle diameters? It seems that it is not. If two (or more) such particles are *simultaneously in excited state*, which lasts significant time in comparison with a period of light oscillations, a resonance interaction or (in quantum interpretation) exchange forces will arise inevitably between them... In consequence of it the emission of both particles should become coherent, connected in phase. Experimentally very strong excitation is required for that (for instance, by a cathode beam) as well as luminescent medium with long duration of molecular “spontaneous” emission (for instance, uranyl salts, which have average duration of spontaneous emission 10^{-3} – 10^{-4} s).”

This book “Microstructure of Light” [26] written almost 70 years ago but still vibrant for the modern optical investigator, is based on Vavilov’s 30-years research (it was written in the end of his life in spite of the busiest period of his administrative activities) and consists of three different parts connected by the concept of “microoptics”. This concept according to the book’s [26] Preface, comprises quantum fluctuations of light at very low powers, interference properties of very small emitters and dependence of this interference on the nature of elementary emitters (dipole, quadrupole etc.), manifestation of excited state lifetimes and polarization properties under pulsed excitation, and influence of the absorbing molecules on the emitting properties of other molecules when both are separated by the distance smaller than the light wavelength. He also included in “microoptics” a light-matter conversion (electron-positron pair creation from gamma-quant). Currently, this book is translated only into German. There is a Fabrikant’s detailed description of the whole book [49], but only in Russian version. To show its topics in more details we are providing here its table of content. It should be noted that each chapter contains several paragraphs which titles we omitted. **Part I.** Experimental investigation of quantum fluctuations of light by visual method: *Chap. 1.* Visual method of measurements of quantum fluctuations; *Chap. 2.* Foreign visual fluctuation measurements; *Chap. 3.* Fluctuations and properties of the eye; *Chap. 4.* Fluctuation measurements of a light flux under low powers. **Part II.** About pre-conditions and some conclusions of elementary study of the interference of light: *Chap. 1.* Limits of realization of optical superposition principle; *Chap. 2.* Bases of the elementary theory of the interference; *Chap. 3.* Interference and the nature of

elementary emitters; *Chap.* 4. Influence of the medium on interference phenomena; *Chap.* 5. Statistical structure of the interference field. **Part III.** Properties of light emitted by the absorptive medium: *Chap.* 1. Emission and absorption of light under inductive connection (see [50]) between molecules; *Chap.* 2. Resonance migration of absorbed energy and depolarization of fluorescence; *Chap.* 3. Concentration quenching of the luminescence and inductive resonance (see [50]); *Chap.* 4. Luminescence quenching and absorption of light.

14.1.3 Method of Observation of Quantum Fluctuations of Light by a Human Eye (Threshold of Human-Eye Sensitivity) and Statistical Structure of Interference Pattern

Another Vavilov's important work is connected with the earlier mentioned experiments on observation of quantum fluctuations of light [26–35] with a human eye as an optical detector. In spite of the results of [38, 39] of 1920 that did not show quantum fluctuations in a feeble light, in 1929 Vavilov and Fedorova started new experiments using feeble light and a human eye accommodated in darkness as a detector [11]. Although some small effect was observed, Vavilov was not satisfied with this work because of possible psychophysiological factors. He did not publish these results.

The series of new experiments by Vavilov's group of 1932–1942 in this direction was triggered by the paper of Barnes and Czerny of 1932 [51] on observation of quantum fluctuations of light by a human eye. Vavilov et al. showed that in spite of the excellent idea and intensity levels in the experiments possible for observation of quantum fluctuations, the results of the experiments of Barnes and Czerny can be explained by physiological fluctuations which were much stronger than quantum fluctuations in the conditions of their experiments. For instance, in paper [51] the eye had no fixed point, so movements of the eye balls were not excluded. Vavilov and Brumberg suggested different from [51] experiment pointing out three necessary conditions for the observation of quantum fluctuations of light by a human eye excluding the influence of the physiological fluctuations: (1) a short-time-exposure (0.1 s); (2) small size of the image on the retina; (3) eye fixation on some point. Figure 14.5 shows the second, 1938–1941 version of their setup with a preservation of the main parts of the earlier setup.

In these experiments, the chin of the observer was fixed, and the observer's eye was fixed on the red spot S . Rotation of a glass plate G and changes of its position permitted to vary the angle between fovea and the place on the retina. The light from the source L was directed to the eye by the mirror m through the aperture O covered with a frosted glass. A disk D with the aperture was rotated using the motor M , so the duration of the light exposure was 0.1 s. Glass filter F , glass plate set P and an absorbing wedge K were used for intensity variation. For absolute

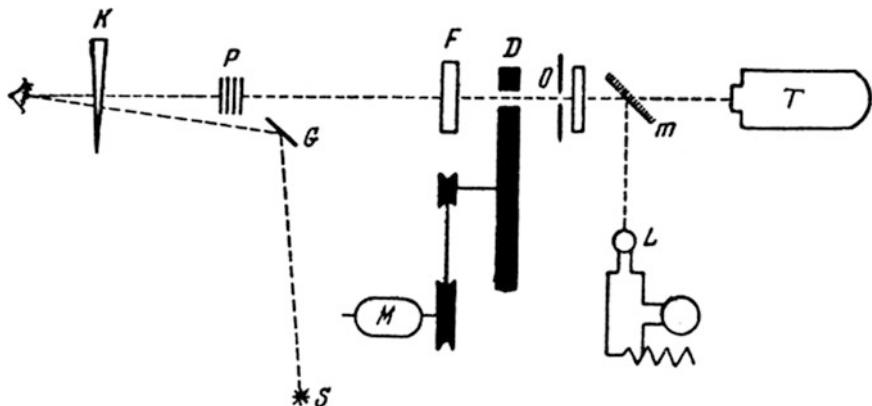


Fig. 14.5 Schematics of Vavilov's second set up for measurements of quantum fluctuations of light by the human eye. See explanations in the text

measurements of the energy of the eye-sensitivity threshold, the mirror m and a frosted glass were removed and the aperture O was irradiated by the “black-body” source T . In this set up the angle dimensions of the round light spot were $6'$. The most of observations were made on the distance of 8° from a fovea. For measurements at different wavelengths, the monochromator with the light source and additional light filters replaced the light source L . The accuracy of energy measurements was $\sim 10\%$. The pupil diameter was defined by the photography of the eye and the scale attached to it. An important part of hundreds series of 1932–1942 year experiments of the Vavilov's group with more than 10 observers, was the statistical method of processing the results. The theory of this method is described in [26–29, 31–35]. References [32–35] in English published in 1943 [32–34] and in 1965 [35] include the results of earlier work of Vavilov's group. Section 14.3 of this chapter contains translation into English of Vavilov's paper [27] with all details of measurements and calculations.

According to Vavilov's group measurements, summarized in [26], the threshold of sensitivity of the retina in the conditions of the experiments for 8 different observers varied from 8 to 47 photons (at 8° from the center of the retina) at 500–550 nm (a number of photons at the cornea in these experiments varied from 108 to 335). Sections 14.3 and 14.5 of the current book contain translation and reprint in English of original Vavilov's papers [27] and [32]. Later work of 1941 by Hecht, Shlaer and Pirenne [52–54], widely cited in the literature, described in details in the Vavilov's book [26], used the similar scheme with insignificant difference (flash duration was 0.001 s). In paper [53] the sensitivity threshold of the human eye retina (at 20° from the center of the retina) at 510 nm was 5–14 photons (with 54 to 108 photons at the cornea). Both Vavilov and Hecht used Poisson distribution of photons. See also Chap. 17 of the current book for Hecht's results and for a brief discussion of modern measurements of the human eye sensitivity threshold.

Vavilov's group measured the eye sensitivity at different spectral ranges. His quantum fluctuation method of evaluation of sensitivity of retina permitted him to observe the second maximum in retina sensitivity in UV region [30] comparable with the first maximum near 510 nm at low light levels (scotopic vision), which was confirmed later by other researchers on the eye without its crystalline lens absorbing UV-light. In connection with human eye sensitivity to light see also Vavilov's popular book "The human eye and the sun" [35], pages 85–106, translated into English.

It should be mentioned Vavilov's remarks in a footnote of his book [26] regarding an incorrect description and interpretation of his group experiments of observation of quantum fluctuations of light by Hecht et al. [53]. See Sect. 14.4 of this chapter with translation of Vavilov's comments and Sect. 17.3 of Chap. 17 with reprint of [53].

14.1.3.1 Independent Fluctuations in Coherent Beams and Statistical Structure of Interference Pattern

Using the human eye as a detector and flashes of faint light at eye's sensitivity threshold (~40–50 "green" photons per flash at the retina in these measurements), Vavilov and Brumberg clearly observed independent intensity fluctuations of two split coherent beams, but sometimes two split beams were seen simultaneously (1933) [26–28]. In [27] the authors made a conclusion: "...the diffraction pattern does not arise immediately everywhere but, rather, emerges gradually, statistically" (see Sect. 14.3.5 with translation of [27]).

A Fresnel double prism was employed to create two coherent beams of a green light [26–27]. In these experiments, two green spots symmetrical to the red fixation point were seen by the eye. At light power close to the sensitivity threshold of the eye, the fluctuations of one spot relative to another were clearly seen. The qualitative measurements showed that these fluctuations in both coherent beams are independent and have a quantum nature. Simultaneous bright spots were also seen sometimes, but these events were rare. The same experiment was made with a Wollaston prism and an unpolarized incident beam [26–27]. Two green spots after a Wollaston prism were polarized orthogonally. The both spots with different polarization had independent fluctuations at light powers near the eye sensitivity threshold. Chapter 18 of this book contains a translation of a Lajos Jánossy paper of 1955 on correlation measurements in a split faint light beams with photomultipliers as detectors and more sophisticated technique.

In [28] untitled "Statistical structure of interference field", Young's double slit experiment was performed for the same purpose and statistical pattern of a double-slit interference in two spots (at maximum and minimum of interference pattern) was compared for a faint light using the human eye as a detector. No photons were detected in a minimum of interference pattern, but in maximum the

spot brightness fluctuated in time. The same experiment was carried out in [28] with a single-slit. *By observation of two different points of the same diffraction maximum it was shown that they fluctuate independently near eye sensitivity threshold.*

14.1.4 Vavilov-Čerenkov Radiation

Another Vavilov's significant contribution to the World science is Vavilov-Čerenkov radiation (in Russian alphabetic order) [36–37, 55–58] known in the West as Čerenkov radiation (sometimes English translators of Soviet/Russian papers do not include Vavilov's name from original papers on this effect in Russian). Seven years after Vavilov's death in 1951, Lebedev Physical Institute scientists were awarded 1958 Nobel Prize in Physics: P. A. Čerenkov who observed this effect, as well as I. M. Frank and I. E. Tamm who explained this effect. Vavilov-Čerenkov radiation was found application in high-energy physics. In this effect the electromagnetic radiation (mostly in visible and near UV spectral region) is emitted when a charged particle gives off energy as it moves through a transparent medium faster than the speed of light in that medium. The radiation is emitted in a cone whose half angle is greater for faster particles and media with higher refractive indices. This is an electromagnetic analog of a hydrodynamic Mach cone.

Tamm in his Nobel lecture noticed that Vavilov played the decisive role in this discovery [37]. See also [8, 9, 11, 36]. Pavel Čerenkov himself wrote about Vavilov's important role in this discovery [58] and that it may have happened only in a strong school of luminescence with developed criteria of distinguishing luminescence from other types of radiations. Earlier Pierre and Marie Curie and other French scientists observed a bluish glow in the dark from glass vessels containing salts of radium, but thought that it was a conventional luminescence. Writing about this, I am not diminishing the role of Pavel A. Čerenkov, who saw this very weak at this time glow that nobody expected to observe. He learned very quickly a technique of measurements, carefully and highly professionally carried out very difficult and challenging measurements in absolute darkness during hours using his own eyes as a sensitive detector. All his measurements were carried out without any errors. The difficulty in this topic and work on it was also enhanced by a skepticism of many colleagues, e.g., famous Russian scientists Ioffe, Kapitsa, Landau, Lebedev Physical Institute co-workers, and even F. Joliot-Curie who visited Moscow at this time [11]. But Vavilov strongly supported this research.

Here is the story of this discovery based on publications of Lebedev Physical Institute scientists where a significant part of this work was carried out (see [8, 9, 36]), although it started in Leningrad. It was important for this discovery that in 1932, Vavilov and Brumberg developed a photometric technique for different spectral regions using the human eye as an instrument for measuring at low light intensities close to the threshold of vision. It is briefly described in the previous

Sect. 14.1.3 of the present chapter. For more details see also [26], Part I, Chap. 4. The visual photometric method of quantum fluctuations of light was very useful at a time when photomultipliers did not yet exist, and it was used in Vavilov's group experiments performed in 1932–1941. This visual technique played a large role in the discovery of the Vavilov-Čerenkov effect.

In 1933 Vavilov proposed to Čerenkov, his Ph.D. student, three possible topics of his thesis and Čerenkov selected one of them: "The luminescence of the uranyl salt solutions under the influence of hard gamma radiation." From beginning of Čerenkov's work Vavilov participated in it. He taught Čerenkov by a technique of measurements in a feeble light using a human eye as a detector, and carried out one or two times per week his own (control) measurements to check his Ph.D. student results staying in a dark room for hours.

In the course of measurements Čerenkov found that a glow is emitted not only by the salt solution but by pure sulfuric acid as well. Vavilov persuaded Čerenkov to continue the experiments with the carefully purified acid. When this did not help, Vavilov proposed checking if other pure solvents also emit light. Čerenkov investigated 16 different solvents of very high purity and found that the pure liquids (e.g., water, parafin, benzol, toluene) all gleamed with nearly equal intensities under the action of gamma rays. He distillated water three times and every time observed the same glow.

Checking if it is a luminescence, Vavilov suggested to quench it by various means, but the mysterious radiation persisted. Vavilov analyzed all of them and arrived at the firm conclusion: "This is not a luminescence, this is a new optical phenomenon not known to science." Vavilov presented a first explanation of this effect—that the new radiation was produced by Compton electrons knocked out from the atoms of the liquid by gamma rays. In 1934 two articles were published in the same volume of the Soviet Physics—Doklady, one signed by Čerenkov [55] with his experimental results, the other by Vavilov in which he correctly postulated fast electrons as the origin of the new phenomenon [56]. Čerenkov offered Vavilov to be a co-author of [55], but Vavilov refused it [8]. He himself called this phenomenon Čerenkov's radiation. It should be noted that Vavilov and Čerenkov also established that the radiation is emitted in a narrow cone close to the direction of the incident gamma-ray flux and is polarized almost along this direction, but not perpendicular to it. A magnetic field deflected the radiation, confirming Vavilov's claim that it originated from charged particles. Vavilov involved in the discussion of this phenomenon two theoreticians Frank and Tamm, who formulated a complete theory of radiation. This priority of Soviet scientists may not be known in the West had Vavilov not advised Čerenkov to submit his manuscript describing this phenomenon to Nature, and when the manuscript was rejected, to send it to Physical Review [57]. For this work all four scientists (Vavilov, Čerenkov, Frank and Tamm) were awarded the highest Soviet scientific prize in 1946, but at a time of the 1958 Nobel prize for this effect Vavilov was not alive.

14.1.5 Vavilov's Contribution to the Luminescence Science and His School of Luminescence Which Created the Basis for First Soviet Lasers

Vavilov gave a strict definition of the luminescence phenomenon, separating it from light reflection and scattering by the emission duration criterion supplementing earlier definition by Wiedemann. Vavilov was among leading researchers in the world who transferred this “insignificant” at this time field into one of the most important directions of science. He was the founder of the Soviet school of luminescence (see Fig. 14.6 on his work in the laboratory). For detailed discussion of his contribution in luminescence science see reviews of his disciples P. P. Feofilov [5, 18], V. L. Levshin [19, 59], and L. V. Levshin [11]. These contributions are as follows:

- (1) The first investigation of the magnitude of the photoluminescence energy output (1924) showing that this output may be large, contrary to the opinion of 1920s that luminescence is infinitesimal in comparison with thermal absorption. Vavilov introduced the concept “fluorescence yield”;
- (2) Formulation the law of constant quantum yield of fluorescence of dye solutions in a wide interval of wavelengths of exciting light. The energy yield increases linearly with a wavelength (Vavilov's law). Vavilov noted that at higher frequencies this law should be violated.
- (3) Establishing fundamental relations for polarized luminescence (jointly with his disciple V. L. Levshin) and its dependence on the molecular symmetry and the nature of elementary emitters (dipoles, quadrupoles etc.);
- (4) Developing theory of luminescence quenching by impurities based on his own experiments. Vavilov's studies have shown (along with those of his disciples A. N. Sevchenko and M. D. Galanin), that depolarization of the luminescence occurs under certain conditions even in solids and in exceedingly viscous

Fig. 14.6 Vavilov in the laboratory (approximately 1949)



solutions in which rotation of the molecules themselves is impossible. Vavilov explained this phenomenon as due to transfer of excitation energy from the system of excited molecules to the unexcited molecules. He derived this theory on the basis of the hypothesis of an inductive interaction between the excited and the unexcited molecules (see [50]). He also obtained quantitative relations for explanation of his experiments on kinetics of fluorescence with quenching impurities. These conceptions of Vavilov have been expanded in the later studies of his disciples M. D. Galanin, B. Ya. Sveshnikov, et al. Galanin has derived a more detailed theory which became one of the basis of more general theory of electron energy transfer in condense media by Förster-Dexter-Galanin.

And to create the first laser in the Soviet Union after Mainman paper, the most prepared groups were groups of Vavilov's luminescence schools at: (1) Lebedev Physical Institute (M. D. Galanin, A. M. Leontovich, Z. A. Chizhikova) and (2) GOI (L. D. Khazov and I. M. Belousova) which independently obtained lasing in ruby. Lebedev Institute group first published their results [60], although GOI group [61] obtained lasing a little bit earlier, but was not published them because of clearance. Both Lebedev Physical Institute and GOI were the main Soviet leaders on high-power laser research.

14.1.6 Vavilov's Works on the History of the Science and Its Popularization

In addition to “The human eye and the sun” [35], Vavilov wrote papers and books about science for general public as well as philosophical and natural science history papers, including biographies of great scientists (volumes III and IV of [6], a total of 1,326 pages). In 1928 his popular book “Experimental basis of theory of relativity” was published. He wrote Newton's biography [62] in harsh years of evacuation in Kazan' during the World War II at the time of the Stalingrad battle. His paper about Newton has been read on his behalf during Newton's Tercentenary celebration by the Royal Society of London in 1946 and published in English in its proceedings [63]. He also sent to London to this event a first ever translation from Latin into “living” language of Newton's *Lectiones Opticae* which he carried out for book of 1946 in Russian [64]. Earlier, in 1927, Vavilov translated from English Newton's *Opticks* [65] from 1721 year edition. Vavilov also left us his diaries 1909–1951 which were published in Russian as a two-volumes book [66]. Reading them, one recognizes his wide interests in art and literature (he admired Goethe) as well as his own poems. Some parts of them are available in Internet [67] as well as the monthly book list on several languages which he read during his last years (1950–1951) [68].

14.1.7 Vavilov's Organizing and Administrative Activities at the State Level

In addition to his scientific achievements, Vavilov was the State-level leader of science who defined the scientific policy of two leading optical institutes (one of them he founded) and was responsible for the scientific policy of the whole country as the President of the Soviet Academy of Sciences. In 1932 he was appointed as the director for research of the State Optical Institute in Leningrad. That same year he was asked by the Academy of Sciences to take charge of the Academy's small Physics Department in Leningrad. Vavilov invited several young physicists to join and organized investigations on the properties of neutrons, radiation-induced luminescence, and colored crystals. In 1934 the Academy of Sciences moved from Leningrad to Moscow, and Vavilov's Physics Department moved there too. Vavilov was determined to turn the small Physics Department into an institute covering the most important fields of physics. He obtained financial support from the government and began to organize the new Physics Institute of the Academy of Sciences. He was appointed as the first director and suggested that the new institute be named after P. N. Lebedev. In Russia, the Lebedev Institute is called "FIAN". He invited distinguished physicists to head the divisions he intended to create and some excellent young physicists to join the staff. So he simultaneously lead both two main Soviet optical institutes: the Lebedev Physical Institute in Moscow and the State Optical Institute in Leningrad [8, 9].

He defined the strategy of the Lebedev Physical Institute as a multidisciplinary institute, attracted the best physicists to it, so the Lebedev Physical Institute quickly became a unique place of scientific spirit and achievements in spite of a harsh time of pushing an ideology into science [69]. And definitely, his scientific school and style permitted Basov and Prokhorov to accomplish their scientific discoveries [70].

In 1945 Stalin offered S. I. Vavilov the position of a President of the Soviet Academy of Sciences. Two and a half years earlier in the prison died his beloved brother Nikolai, a world-famous geneticist and plant breeder, the author of the law of genetic homology and a creator in Leningrad of the world's largest collection of plant seeds [10, 71] (see, e.g., [72] about destruction of genetics in the USSR at this time). Sergei Vavilov knew that this status will force him to tell terrible ritual words and make disgusting compromises, but nobody at this time was able to reject Stalin's offering. He also knew that should he escape the presidency, Stalin would appoint one of his favorites who would ruin Soviet science completely. Documents of this time about other candidates confirm this. Some authors who did not know Vavilov, criticized him for the acceptance of this position, see, e.g., book [73] of 1970 Nobel prize winner in literature A. I. Solzhenitsyn and [74–75] published recently in English by the author of younger generation. 2003 Nobel Prize winner in physics V. L. Ginzburg knowing Vavilov well, convinced, that all accusations against S. I. Vavilov are absolutely inconsistent [76], wrote a letter to Solzhenitsyn to inform him (with the enclosure of the corresponding materials) that his appraisal of Vavilov in "The Gulag Archipelago" was wrong. Solzhenitsyn called him in

response to express his satisfaction with the elucidation of the truth [76]. Another Nobel prize winner, I. M. Frank wrote in [3], page 47: “Many are now asking how Vavilov could agree to become President even though his beloved brother had perished in prison. A pertinent question to those asking is what would have happened if he had refused? I am not sure that he would have been allowed to stay alive as had happened with Kapitsa who had shown obstinacy. ...Even if Stalin had not destroyed him he would definitely have been stripped of his rank and removed from all his positions, in particular, from running his favorite creation, the Lebedev Physical Institute of the Soviet Academy of Sciences. I am absolutely convinced that the least of Vavilov’s concerns was his own destiny. He felt deeply his personal responsibility for the destiny of science and culture. I am sure that if Vavilov had been purged, the Lebedev Physical Institute would have been branded an asylum for enemies of the people. It was a well-known fact that we, his absolutely devoted disciples in science and life, were greatly indebted to him. He would not have been able to protect us from inevitable persecution.”

S. E. Frish, who worked at the State Optical Institute at this time in his sincere book [77] criticizes Vavilov for his style to “soften the rough angles” in human relations and compromises in differences to Frish’s teacher D. S. Rozhdestvensky, the founder of State Optical Institute. He also blames Vavilov that he was not able to prevent Rozhdestvensky’s own leaving of this institution to the Leningrad State University because of closing his scientific direction by the administrative director. It seems that Vavilov can not do it at this time. But without Vavilov’s help and support in difficult years we would not know Fabrikant’s achievements [78] (Fabrikant as son of politically sentenced father should leave Moscow State University). Vavilov also helped Kapitsa when Kapitsa was dismissed from his institution, although he knew that Kapitsa did not appreciate his scientific work (luminescence as a scientific field and discovery of Vavilov-Čerenkov radiation), as well as his personality.

The value of this great person one can also recognize from many recollections of people who worked with him, knew him well and left many warm, sincere and informal notes about him [3, 7–18]. Nobel prize winner I. M. Frank was happy when he finished his moral duty which he proposed for himself in preparing for publication the book [3] about his teacher. Vavilov was a kind, responsive person as all people who new him state and he never suppressed other people. In the 1930s and 1940s Vavilov did his best to help the victims of Stalin’s terror. He wrote letters addressed to supreme rulers in defense of arrested scientists, not even being personally familiar with them, and helped them. His salary was mainly used for sending money to the people in need [7].

In the President position (see Fig. 14.7), Vavilov provided gigantic on the scale, enormous support of the Soviet science protecting the whole directions of it from destruction (when he has been able to do it) and helping other people, although he was not able to save Soviet genetics from destruction [72]. One example of this time event is the All-Union conference on the philosophical aspects of physics that was planned for 1949 and that could have had disastrous consequences for Soviet physics. The main targets were quantum mechanics and theory of relativity as

Fig. 14.7 Vavilov shortly after his election as a President of the Soviet Academy of Sciences (SAS) in the office of the Presidium (Headquarters) of the SAS (1945)



“idealistic” and “cosmopolitan” theories, see, e.g., the book [79], pp. 208–211. According to L. A. Artsimovich (the leader of the Soviet Tokamak project) who spoke after Stalin’s death with Beria (the chief of security and secret police and the chair of the special committee on the atomic bomb), three leading physicists asked Beria to call off this conference that will interfere with the atomic project [79]. But the author of [74, 75] who investigated archives of this time, did not find any documents why this conference was called off [80]. It was known that S. I. Vavilov did all what he had been able to do to delay this devastating conference, and if it would be in his forces, to cancel it. For instance, one day in 1948, when his co-worker asked him about some minor administrative issue, not recognizing Vavilov’s preoccupation with something else, Vavilov said sadly: “Alas, Vsevolod Vasilievich, I have to be saving Soviet physics now” ([3], p. 208).

All his efforts and compromises connected with his President position, required from him physical forces beyond his health limit. I. M. Frank wrote (from [3]): “Already being a President of the Soviet Academy of Sciences, he told me “Every time, when I was called to visit the Kremlin, I am not sure, whether I will return home or they will take me to Lubyanka (the headquarters of the secret police)”. He worked so hard that his colleagues from the State Optical Institute wrote a letter to the government expressing concerns about his health. Hence after the heart attack (nine scars on his heart were found after his death) he untimely deceased in 1951 on

the eve of his sixtieth birthday. For two days, in spite of a strong frost, people stayed in a long line to pay last respect to him.

14.1.8 Some Biographical Information About S. I. Vavilov and His Family

At the end of this essay I include some brief biographical information about Sergei Vavilov. He was born in Moscow in 1891 and graduated with honor from Moscow State University. His father was a prosperous textile merchant who made himself starting from a farmer's work (he was the son of the serf-farmer who was liberated by a 1861 year reform) and from the age 12 working as a boy in a store of a big Moscow factory. At the age of 15 he was already a director of one of the factory's stores. He gave a good education to his two sons and hoped they would inherit and continue his business. However, Sergei and his elder brother Nikolai both decided to become scientists. His mother was a daughter of a textile artist of the same factory and devoted her life to her children. As a second-year student Sergei Vavilov started work in the laboratory of P. N. Lebedev (who first measured the pressure of light), where he carried out his first research paper which was awarded by one of the societies belong to Moscow State University. Vavilov became a great admirer of Lebedev's style of work—an interest in the fundamental problems of physics blended with careful experiments.

In 1914–1918, during World War I he served in the army in combat among radio engineering troops where he wrote several scientific papers on Radioengineering. Figures 14.8 and 14.9 show his young-year photographs. When by the end of 1917 the Russian front had collapsed because of the Revolution, he was taken prisoner but escaped and in February 1918 appeared in Moscow.

In 1918 Vavilov started to work in the Physics and Biophysics Institute headed by P. P. Lazarev, a disciple of Lebedev who proposed that Vavilov pursue the topic of physical optics. In this institute he worked until 1930 and simultaneously taught physics at Moscow State University (1918–1932). He got married in 1920 to O. M. Bagrinovskaya with whom he happily shared the whole life. His son V. S. Vavilov became later a well-known scientist in the field of physics of semiconductors. In 1926 S. I. Vavilov worked several months in Berlin's laboratory of P. Pringsheim. During this stay in Germany, Vavilov met many outstanding physicists including A. Einstein, M. Born, M. Laue, J. Franck, O. Hahn, and L. Meitner. From 1932 he was director and organizer of P. N. Lebedev Physical Institute (Moscow), and simultaneously held the position of director in research of the State Optical Institute (Leningrad) in 1932–1945. In 1945 he was elected President of the Soviet Academy of Sciences. He died in 1951.

For full bibliography of Vavilov's publications see [81–82]. Reference [6] contains his publications in four volumes.



Fig. 14.8 Vavilov in his young years during the World War I in 1914 or 1915



Fig. 14.9 Sergei Vavilov (right) in his young years with his brother Nikolaj (left) and his mother Alexander M. Vavilova (center) at the time of his leave from the World War I front (December of 1916)

Acknowledgements

The author thanks A. M. Leontovich, B. M. Bolotovskii, Yu. N. Vavilov, E. B. Alexandrov, V. N. Zadkov for valuable comments after reading this chapter. Vavilov's photographs were published with permission of the office of history of the P. N. Lebedev Physical Institute, Russian Academy of Sciences (RAS), and the S. I. Vavilov Institute for the History of Science and Technology of RAS (Laboratory of Scientific and Applied Photography, Cinematography and Television LAFOKI). Some citations were used with permission from the Russian journal *Uspekhi Fizicheskikh Nauk* (Physics-*Uspekhi*).

References

1. R.G.W. Brown and E.R. Pike: A history of optical and optoelectronic physics in the twentieth century, in the book *Twentieth Century Physics*, L.M. Brown, A. Pais, B. Pippard, eds. (Inst. of Physics/American Inst. Phys. Press, Bristol (England) and Philadelphia//NY, 1995), Vol. III, Chap. 18, p 1403.
2. J.D. Bernal: Academician S.I. Vavilov, *Nature*, No 4277, 679, October 20 (1951).
3. *Sergei Ivanovich Vavilov: Essays and Recollections*, I.M. Frank, Ed. (3d edition, Nauka Publ., Moscow, 1991), in Russian.
4. O.N. Krokhin: S.I. Vavilov—the founder of the P.N. Lebedev Physics Institute, *Physics-Uspekhi* **44** (10), 1026–1028 (2001).
5. P.P. Feofilov: Sergei Ivanovich Vavilov, in *50 years of S.I. Vavilov's State Optical Institute (1918–1968)*, Yu.N. Gorokhovsky, ed., Mashinostroyeniye Publ., Leningrad, 1968, in Russian.
6. S.I. Vavilov, *Collected Papers (Sobranie sochinenii)*, Volumes I—IV (USSR Academy of Sciences Publ., Moscow, 1952–1956), in Russian.
7. E.L. Feinberg, *Physicists: Epoch and Personalities. History of Modern Physical Sciences*, Vol. 4, World Scient. Publ., NJ, 2011.
8. B.M. Bolotovski, Y.N. Vavilov, A.P. Shmeleva: Sergei Vavilov: luminary of Russian physics, *CERN Courier*, November 12, 2004. See website <http://cerncourier.com/cws/article/cern/29202>.
9. B.M. Bolotovskii, Yu.N. Vavilov and A.N. Kirkin: Sergei Ivanovich Vavilov—the man and the scientist: a view from the threshold of the 21st century, *Physics-Uspekhi*, **41**, N5, 487 (1998). Translated from Uspekhi Fizicheskikh Nauk, **168**, N5, 551–570 (1998).
10. Yu.N. Vavilov, *In a long search. Book about brothers Nikolai and Sergei Vavilov* (FIAN Publ., Moscow, 2004), in Russian.
11. L.V. Levshin, *Sergei Ivanovich Vavilov* (Nauka publ., Moscow, 2003), in Russian.
12. *To the history of FIAN (P.N. Lebedev Physical Institute), Portrait ser.*, is. 2, part 1, *Sergei Ivanovich Vavilov. New features to the portrait*, Ed. B.M. Berezanskaya (FIAN Publ., Moscow, 2004), in Russian. See also website <http://ellphi.lebedev.ru/12/SIVavil.pdf>.
13. S.I. Vavilov Institute for the History of Science and Technology of the Russian Academy of Sciences, Project “Sergei Ivanovich Vavilov”, *Bibliography and Recollections about S.I. Vavilov*, <http://www2.ihst.ru/projects/vavilov/index.htm>, in Russian.
14. M.V. Alfimov, B.M. Bolotovskii, Yu.S. Osipov, V.L. Ginzburg, O.N. Krokhin, E.L. Feinberg, A.M. Bonch-Bruevich: Reports about S.I. Vavilov at the special meeting of the Presidium of the Russian Academy of Sciences, the Bureau of the General Physics and Astronomy Division, RAS and the Scientific Council of the P.N. Lebedev Physics Institute, RAS on the occasion of the 110th anniversary of the birth of Sergei Ivanovich Vavilov, and XXV Jubilee Vavilov

- Reading (28 March 2001), *Physics-Uspekhi*, **44**, N 10, 1017–1036 (2001). Translated from *Uspekhi Phyzicheskikh Nauk*, **171**, N 10, 1071–1090 (2001).
- 15. A.A. Lebedev: Fragments of recollections about S.I. Vavilov, in Russian, see website <http://www2.ihst.ru/projects/vavilov/memory/lebedev.htm>.
 - 16. B.M. Bolotowsky, Yu.N. Vavilov: Sergei Vavilov, honored in Russia, still little known in the West”, *Phys. Today*, **11**, 13, December (1995).
 - 17. B.P. Keler, *Sergei Vavilov*, Moscow (1975), in Russian.
 - 18. P.P. Feofilov: S.I. Vavilov, the founder of the Soviet school of luminescence, *Soviet Phys. Uspekhi*, **4**, N 5, 770–775 (1982), translated from *Usp. Fiz. Nauk*, **75**, N 10, 277–286 (1961).
 - 19. V.L. Levshin: The development of S.I. Vavilov’s ideas in the field of luminescence, *Sov. Phys. – Uspekhi*, **4**, N 5, 747–753 (1962). Translated from *Usp. Fiz. Nauk*, **75**, N 10, 241–250 (1961).
 - 20. *President of the USSR Academy of Sciences S.I. Vavilov—the 120 anniversary of his birthday*, CD, Presidium of the Russian Academy of Sciences Publ. (2011).
 - 21. R. Pike and R.G.W. Brown: More on Vavilov’s contributions to 20th—century Physics, *Phys. Today*, 120 and 122, September (1996).
 - 22. A.V. Masalov and Z.A. Chizhikova: S.I. Vavilov and nonlinear optics, *Phys.-Uspekhi*, **54**, N 12. Translated from *Usp. Fiz. Nauk* **181**, 1329–1334 (2011)
 - 23. N.I. Zheludev: Single nanoparticle as photonic switch and optical memory element, *J. Opt. A: Pure Appl. Opt.*, **8**, S1–S8 (2006).
 - 24. S. Frish: The development of Soviet Optics and Spectroscopy during the past fifty years, *Appl. Opt.*, **6**, N 11, 1783–1792 (1967).
 - 25. S.J. Wavilow and W.L. Lewschin: Die Beziehungen zwischen Fluoreszenz und Phosphoreszenz in festen und flüssigen Medien, *Zeitschrift für Physik*, 31–35, 920–936 (1926), in German. Translated into Russian in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 195–207.
 - 26. S.I. Vavilov, *Mikrostruktura Sveta (Issledovaniya i Ocherki)* [Microstructure of Light (Investigations and Essays)], (USSR Academy of Science Publishing, Moscow, 1950), 198 p., in Russian. See also German translation S.I. Wawilow, *Die Microstruktur des Lichtes* (Akademie-Verlag, Berlin, 1954).
 - 27. E. Brumberg and S. Vavilov (S. Wawilow): Visuelle Messungen der statistischen Photonenschwankungen (Visual measurements of statistical fluctuations of photons), *Izvestia Acad. Nauk SSSR, Ser. Math. (Bulletin de l’Académie des Sciences de l’URSS)*, OMEN, N 7, 919–941 (1933), in German. Reprinted in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 345–364, in Russian.
 - 28. E.M. Brumberg and S.I. Vavilov: Statistical structure of interference field, *Dokl. Akad. Nauk SSSR (Soviet Physics: Doklady; Comptes Rendus de l’Académie des Sciences de l’URSS)* **3 (III)**, N 5, 322–325 (in Russian), 325–328 (in German) (1934).

29. S.I. Vavilov: Directions of development of the Optical Institute, *Izvestia Acad. Nauk SSSR*, (*Bulletin de l'Académie des Sciences de l'URSS*), Ser. Fiz. No. 1–2, 163–188 (1936), in Russian. See also the same paper in *Uspekhi Fizicheskikh Nauk* (Physics-Uspekhi) **16**, 872–896 (1936), in Russian.
30. S.I. Wawilow: Sensitivity of the retina to the ultra-violet spectrum, *Dokl. Acad. Nauk USSR*, (*Soviet Physics: Doklady; Comptes Rendus de l' Académie des Sciences de l'URSS*), **21**, N 8, 373–375 (1938), in English.
31. S.I. Vavilov: Experimental investigation of light quantum fluctuations by visual method, *Uspekhi Fizicheskikh Nauk*, **36**, 247–283 (1948), in Russian.
32. E.M. Brumberg, S.I. Vavilov and Z.M. Sverdlov: Visual measurements of quantum fluctuations. I. The threshold of Vision as compared with the results of fluctuation measurements, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), Vol. VII, N 1, 1–8 (1943).
33. S.I. Vavilov and T.V. Timofeeva: Visual measurements of quantum fluctuations. II. Fluctuations when the eye is light-adapted, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), Vol. VII, N 1, 9–11 (1943).
34. S.I. Vavilov and T.V. Timofeeva: Visual measurements of quantum fluctuations. III. The dependence of the visual fluctuations on the wave-length, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), Vol. VII, 12–17 (1943).
35. S.I. Vavilov, *The Human Eye and the Sun, “Hot” and “Cold” Light*, Pergamon Press (1965), 156 pages. Translated from Russian, see e.g., last edition *The Eye and the Sun* (Amfora Publ., St-Petersburg, 2006).
36. B.M. Bolotovskii: Vavilov-Čerenkov effect: history of discovery, modern state of the problem, pp. 114–133, in *To the history of FIAN (P.N. Lebedev Physical Institute), Portrait ser.*, is. 2, part 1, *Sergei Ivanovich Vavilov. New features to the portrait*, Ed. B.M. Berezanskaya (FIAN Publ., Moscow, 2004), in Russian. See also website <http://ellphi.lebedev.ru/12/SIVavil.pdf>.
37. I.E. Tamm, *Nobel Lecture*, p. 3 (472), December 11 (1958). See website http://nobelprize.org/nobel_prizes/physics/laureates/1958/tamm-lecture.pdf.
38. S.I. Vavilov: About the independence of the light absorption coefficient from brightness, *Izvestia*, Physical Institute under Moscow Scientific Institute, **1**, is. III, 92 (1920), in Russian. Reprinted in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 80–83.
39. S.I. Vavilov: Light absorption at very-low intensities, *Izvestia*, Physical Institute under Moscow Scientific Institute, **1**, is. III, 96 (1920), in Russian. Reprinted in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 84–87.
40. In his book [26], p.69, Vavilov proves that his experiments at intensities close to the sensitivity threshold of human-eye at 550 nm (lower than $1 \frac{\text{photon}}{\text{cm}^2\text{s}}$) showed that the classical character of absorption suggested by Planck in his book *Wärmestrahlung* (2 Aufl., 1913) contradicts with his suggestion of quantum nature of photon emission. At 5.6 times diminishing transmission by

one of his samples doped with a Rhodamin B dye, it would take $\sim 10^{15}$ s at incident intensity $< 3 \cdot 10^{-12} \frac{\text{erg}}{\text{cm}^2\text{s}}$ for each molecule to receive one quant energy under classical, continuous absorption. In this case, molecules should not emit light and absorption coefficients at very low intensities should be very high.

41. A. König, *Gesammelte Abhandlungen zur Physiologischen Optik* (Leipzig, 1903), pp. 116–139.
42. I.G. Priest, H.J. McNicholas, and M.K. Frehafer: Some tests of the precision and reliability of measurements of spectral transmission by the Koenig-Martens spectrophotometer, *JOSA*, **8**, is.1, 201–211 (1924).
43. G.N. Lewis, D. Lipkin, T.T. Magel: Reversible photochemical processes in rigid media. A study of the phosphorescent state, *J. Am. Chem. Soc.* **63**, 3005–3018 (1941).
44. R.W. Boyd: History of research in nonlinear optics at the Institute of Optics, in *A Jewel in the Crown, 75th Anniversary Essays, the Institute of Optics, University of Rochester*, Ed. By C.R. Stroud, Jr. (Meliora Press, Rochester, NY, 2004).
45. M.A. Kramer, W.R. Tompkin, R.W. Boyd: Nonlinear optical interactions in fluorescein-doped boric acid glass, *Phys. Rev. A* **34**, 2026 (1986).
46. B. O'Brian: A brightness limiting phosphor telescope with selective action on the brighter portion of an image, *JOSA*, **36**, is. 12, 709 (1946).
47. R.W. Boyd, S.G. Lukishova, Y.-R. Shen, Eds, *Self-focusing: Past and Present. Fundamentals and Prospects*, Springer Series: Topics in Applied Physics, Vol. 114, Springer NY (2009).
48. According to the superposition principle, a result from several independent actions is the summation of the effects caused by each action separately. In part II, Chapter 1 “Limits of validation of superposition principle” of book [26], Vavilov wrote: “...in spite of a “linearity” of optics of a material medium,... it is possible nevertheless to show in the general case, that such linearity, and therefore, and superposition under light propagation in material medium must be violated. This is a consequence of quantum properties of light and matter”.
49. V.A. Fabrikant: S.I. Vavilov’s book “Microstructure of light”, *Uspekhi Fizicheskikh Nauk*, **43**, is. 5, 117–134 (1951), in Russian.
50. Vavilov uses the terms “inductive connection” and “inductive resonance”. As described in his book [26], p. 183, he uses concepts of F. Perrin published in C. R., **184**, 1097 (1927) on “molecular induction”, which means luminescence quenching by molecules of the other sort. Vavilov and his co-author published paper I.M. Frank u. S.I. Wawilow, Z. Phys. **69**, 100 (1931), studying experimentally this effect (see also Fig. 58, p. 186 of [26]). First Vavilov’s experiment on it was reported in his paper S.I. Wawilow, Z. Phys. **53**, 665 (1929). In the book [26] Vavilov provided reference on similar later works by Th. Förster, Z. Elektrochem., **53**, 93 (1949) and Z. Naturforsch., **4a**, 321 (1949).

51. R.B. Barnes and M. Czerny: Läßt sich ein Schroteffekt der Photonen mit dem Auge beobachten? *Z. Phys.* **79**, 436–449 (1932).
52. S. Hecht, S. Shlaer and M.H. Pirenne: Energy at the threshold of vision, *Science* **20**, 585–587 (1941).
53. S. Hecht, S. Shlaer and M.H. Pirenne: Energy, quanta and vision, *J. Gen. Physiol.*, **25**, 819–840 (1942).
54. S. Hecht, The quantum relations of vision, *JOSA* **32**, 42–49 (1942).
55. P.A. Čerenkov: Visible radiation of pure liquids under the action of γ -radiation, *Dokl. Akad. Nauk USSR (Soviet Physics: Doklady; Comptes Rendus de l' Académie des Sciences de l'URSS)*, **2**, № 8, 451–457 (1934), in Russian and German.
56. S.I. Vavilov: About possible explanation of a blue γ -glow of liquids, *Dokl. Akad. Nauk USSR (Soviet Physics: Doklady; Comptes Rendus de l' Académie des Sciences de l'URSS)*, **2**, № 8, 457–461 (1934), in Russian and German.
57. P.A. Čerenkov: Visible radiation produced by electrons moving in a medium with velocities exceeding that of light, *Phys. Rev.*, **52**, 378–379 (1937).
58. P.A. Čerenkov: Serving science, in *Sergei Ivanovich Vavilov. Essays and Recollections, I. Frank, ed., (3d Edition., Nauka, Moscow, 1991)*, pp. 217–220, in Russian. See also website <http://www.ihst.ru/projects/vavilov/index.htm>.
59. V.L. Levshin, A.N. Terenin, I.M. Frank: The development of S.I. Vavilov's work in physics", *Sov. Phys. Usp.* **4**, N 5, 730–736 (1962), translated from *Usp. Fiz. Nauk* **75**, 215–225 (1961).
60. A.M. Leontovich, Z.A. Chizikova: On the creation of the first ruby laser in Moscow, *Phys. Usp.* **54**, 77–85 (2011). Translated from *Usp. Phys. Nauk* **181**, 82–91 (2011).
61. I.M. Belousova: The laser in the USSR: the first steps, *Phys. Usp.* **54**, 73–75 (2011). Translated from *Usp. Phys. Nauk* **181**, 79–81 (2011).
62. S.I. Vavilov, *Isaak Newton (1643–1721)*, (USSR Academy of Sciences Publ., Moscow-Leningrad, 1943), in Russian.
63. S.I. Vavilov: Newton and the atomic theory, in *Newton Tercentenary celebrations*, 15–19 July 1946, (The Royal Society, Cambridge Univ. Press, 1947), pp. 43–55.
64. I. Newton, *Lectiones Opticae* (The Optical Lectures, 1670–1672). Russian translation from Latin by S.I. Vavilov (USSR Academy of Sciences Publ., Moscow, 1946), 295 p. including paper of S.I. Vavilov "Lectures on Optics by I. Newton", pp. 260–275. See translation from Russian book into German: (Wien. Neues Österreich, 1948, 175 S).
65. I. Newton, *Opticks or a treatise of the reflections, refractions, inflections and colours of light* (London, 3d edition, 1721). Russian translation from English by S.I. Vavilov (State Publ., Moscow-Leningrad, 1927), 371 p.
66. S.I. Vavilov, *Diaries of 1909–1951*, two volumes, Moscow, Nauka Publ. (2016), in Russian.
67. S.I. Vavilov, *Diaries of 1939–1951*, <http://vivovoco.astronet.ru/VV/JOURNAL/VIET/VAVIDIARY.HTM#00>, in Russian.

68. Books read by S.I. Vavilov in 1950–1951, <http://vivovoco.astronet.ru/VV/JOURNAL/VIET/VAIDIARY.HTM#000>, in Russian.
69. A.G. Vitukhnovsky: Vavilov and FIAN—view from the year 2016, *Uspekhi Phyz. Nauk* **186**, N12, 1360–1367 (2016).
70. N.V. Karlov, O.N. Krokhin, and S.G. Lukishova: History of quantum electronics at the Moscow Lebedev and General Physics Institutes: Nikolaj Basov and Alexander Prokhorov, *Appl. Opt.*, Special LaserFest issue, **49**, No. 25, F32–F46 (2010).
71. N.I. Vavilov, *Origin and Geography of Cultivated Plants*, Cambridge Univ. Press (1992).
72. V.N. Soyfer, *Scientific Fraud and Deception under Stalin's Reign*, in submission.
73. A.I. Solzhenitsyn, *Gulag Arkipelago, 1918–1956: an experiment in literary investigation* (Harper & Row, NY, 1973).
74. A.B. Kojevnikov, *Stalin's Great Science. The Times and Adventures of Soviet Physicists*, History of Modern Physical Sciences, Vol. 2 (Imperial College Press, London, 2004), 360 p.
75. A. Kojevnikov: President of Stalin's Academy: the mask and responsibility of Sergei Vavilov, *Isis*, the Univ. of Chicago Press, **87**, N 1, 18–50 (1996).
76. V.L. Ginzburg: About Sergei Ivanovich Vavilov, in *About Science, Myself and Others*, Inst. of Phys. Publ., Bristol and Philadelphia, 2005, Part II, Chapter **16**, pp. 310–317. See also *Physics-Uspekhi*, **44** (10), 1024–1026 (2001). Translated from *Uspekhi Phyzicheskikh Nauk*, **171**, N 10, 1077–1080 (2001).
77. S.E. Frish, *Through the Prism of the Time*, Moscow, Political Literature Publ. (1992), in Russian.
78. S.G. Lukishova: Valentin A. Fabrikant: negative absorption, his 1951 patent application for amplification of electromagnetic radiation (ultraviolet, visible, infrared and radio spectral regions) and his experiments, *J. Europ. Opt. Soc.*, Special issue devoted to the Laser Anniversary, **5**, 10045S-1-10 (2010).
79. D. Halloway, *Stalin and the Bomb, The Soviet Union and Atomic Energy (1939–1956)*. (Yale University Press, New Haven & London, 1994), 464 p.
80. A.B. Kozhevnikov, Private communication, 2016.
81. Bibliography of S.I. Vavilov's publications, *Uspekhi Fiz. Nauk* **XLIV**, is. 1, 136–172, (1951), in Russian.
82. *Materials to bibliographies of USSR scientists*, Physics Ser., issue 22, *S.I. Vavilov (1891–1951)*, V.L. Levshin and T.O. Vreden-Kobetskaya, third edition, Moscow (1979).

14.2 The First Observation of Nonlinear Optical Effect (Saturation of Absorption): Translation of Excerpts from an Original Paper of S. I. Vavilov and V. L. Levshin “The Relation Between Fluorescence and Phosphorescence in Solid and Liquid Media”

Translated into English by Vladimir Rogovoi from Russian book S. I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 195–207. See also S. J. Wavilow and W. L. Lewschin: Die Beziehungen zwischen Fluoreszenz und Phosphoreszenz in festen und flüssigen Medien, *Zeitschrift für Physik*, **31–35**, 920–936 (1926), in German.

This paper was presented by Peter Pringsheim on 12 May 1926 at the Berlin Colloquium during Vavilov’s visit of a Pringsheim’s laboratory in Berlin. On May 17 1926 Vavilov presented this work in Göttingen.

Permission for publishing and translation from Nauka Publisher, Moscow, Russia.

Abstract (1) Two new phosphoroscopes for investigation of phosphorescence of very short duration (from 10^{-6} to 10^{-4} s) have been described. (2) Observations made with these instruments have led us to conclude that fluorescence and phosphorescence in solids and very viscous liquids correspond to two substantially different processes between which there is no gradual transition, although optical properties (spectrum, polarization) may coincide. (3) Only phosphorescence has been observed in uranium glass, while fluorescence has been completely absent. (4) A method has been proposed for detecting short times of the excited state of phosphorescent molecules based on the absorption decay upon illumination with intense light of a spark.

14.2.1 *The Possibility of Reducing the Absorption of Fluorescent and Phosphorescent Objects upon Their Illumination with Intense Light of a Spark (§ 4 of the Original Paper)*

The applicability limits of the Bouguer law on the independence of the absorption ability of an object on the intensity of the incident light are very wide. A few years ago, one of us [S. I. Vavilov, *Izvest. Phys. Inst. (Moscow)* **1**, 92 (1920)] has succeeded in showing that the absorption coefficient of a medium, such as a gelatin plate colored with rhodamine, remains unchanged upon varying the energy of the incident visible light within the range roughly between 10^{-10} and 10^8 erg/s per squared centimeter. From the point of view of quantum theory, one can expect departures from this law in the range of very low energy fluxes and upon illumination with extremely intense light.

At vanishingly low intensities, in accordance with quantum theory, variations of the absorption in time should occur; a quantum of light should be either totally absorbed by the medium or transmitted by it as a whole. In accordance with quantum theory, deviations at high intensities are related to the finite duration of excited states of molecules. The Bouguer law can be expressed in the following form:

$$I = I_0 \exp^{-N\alpha}, \quad (14.1)$$

where I_0 is the incident energy, I is the transmitted energy of monochromatic light, and α is the average absorption ability per one molecule. Upon prolonged illumination, a certain part x of the total number N of absorbing molecules will be in the excited state, and, in accordance with this, expression (14.1) should be replaced by

$$I = I_0 \exp^{-N(1-x)\alpha}. \quad (14.2)$$

In the latter expression $x = f(I_0)$, and the Bouguer law ceases to be rigorously fulfilled. In the case of small values of x , the absorbed energy I_{abs} can be determined from expression (14.1), so that

$$x = \frac{I_{\text{abs}}\tau}{N \cdot h\nu}, \quad (14.3)$$

where τ is the average duration of the excited state, and $h\nu$ is the energy of the absorbed light quantum.

Measurement of energy of concentrated arc light that was absorbed in the visible part of the spectrum by fluorescein solutions showed that in the arc light only those deviations from (14.1) can be noticeable that correspond to τ at the order of magnitude not smaller than 10^{-4} s. For this reason, it seems hardly possible to observe noticeable deviations from the Bouguer law in liquid solutions using concentrated sunlight or arc light. For these same reasons, the proportionality between the intensity of the incident light and the excited fluorescence in liquid solutions should be rigorously fulfilled over a wide range of (14.1). At the same time, in phosphors with long lasting phosphorescence, the above-noted decrease in the absorption manifests itself in the well-known phenomenon of the phosphorescence intensity rise and saturation of phosphors upon their prolonged illumination.

For qualitative observation of this effect, conditions in which light from a spark is used seem to be more favorable. In this case, a very intense illumination of short duration (from 10^{-5} to 10^{-8} s) can be obtained. Let us assume that $T \leq \tau$. If we assume for simplicity that, within the time T , the average value of the intensity of spark light remains unchanged, then we can easily show that, upon monochromatic illumination, the average value of x can be expressed as follows:

$$x = \frac{E}{2hv \cdot N}, \quad (14.4)$$

where E is the energy that was absorbed within the spark emission time.

Tentative measurements of the spark emission energy were conducted. Two Leyden jars were charged from a transformer. A discharge (to oscillations of which, a wavelength of about 50 m corresponded) was initiated between aluminum electrodes, the spacing between which was 10–12 mm (in other experiments, copper, zinc, and cadmium electrodes were also used). The spark light was concentrated into a spot with an area of 0.16 cm^2 . This concentrated light beam was transmitted through a glass cell with a thickness of 2 mm, which was filled with a solution of fluorescein in glycerol with a concentration of 10^{-4} g/cm^3 . Energy measurements were performed using a thermoelement, a galvanometer, and a Hefner candle as a reference. It turned out that the energy of the visible radiation that was absorbed in the area onto which the light was concentrated during the discharge time was 900 erg. Substituting $hv = 4 \times 10^{-12} \text{ erg}$ (maximum of absorption of the fluorescein solution) into expression (14.4) and assuming that $N = 5.8 \times 10^{15}$ and $E = 9 \times 10^2 \text{ erg}$, we find that x is roughly equal to 2%. If the condition $T \leq \tau$ is implemented, then x remains unchanged for all values of τ that satisfy this condition. In our experiments, the discharge was not aperiodic, and its damping decrement was about 0.1–0.2. The average duration of the discharge was therefore on the order of 10^{-6} s .

In numerous absorption measurements, the concentrated spark light was transmitted through a cell with a thickness of 2 mm. One half of the image of the spark passed through the solution, while the other half passed through glass. The corresponding halves of the image were directed onto the two slits of a spectrophotometer. A blue or a yellow light filter, which reduced the intensity of the exciting light in the examined range of the spectrum by a factor of five to ten, was initially placed between the light source and the cell, and then between the cell and the spectrophotometer. To achieve a high accuracy, it was necessary to perform many measurements (from 50 to 100 for either position of the light filter). The latter measurements, which were performed especially carefully, led us to a negative result. Upon using the light filter, i.e., upon a tenfold change in the incident energy, the absorption ability of the solutions of fluorescein in water and glycerol remained unchanged with an accuracy of up to 0.3%. Hence, in accordance with the experiments that were described above, we can conclude that in this case $\tau < 10^{-6} \text{ s}$.

Positive results were obtained with uranium glass. This experiment was performed in the same way, but with the difference that, instead of the cell with the solution, a uranium glass plate was installed through which one half of the image of the spark passed, while the other half of the image was incident immediately onto the slit of the spectrophotometer. The light intensity was reduced using yellow glass, which decreased the intensity of light at a wavelength of $454 \text{ m}\mu$ by a factor of three. At either position of the light filter, 50 measurements were made. The obtained average values of the rotation angle of the ocular Nicol prism of the spectrophotometer are $58.7^\circ \pm 0.1$ and $58.3^\circ \pm 0.1$.

This corresponds to the absorption coefficients equal to 2.576 and 2.544. The first value was obtained when the light filter was placed in front of the cell, while the second value is for the light filter behind the cell. This corresponds to a decrease in the absorption coefficient by approximately 1.5%. The observed positive effect only slightly exceeds the achieved degree of accuracy, but it is consistent with that which could be expected in uranium glass. In this case, $\tau \gg 10^{-6}$. It should be noted that, in the case of the concentrated arc light, we also detected noticeable traces of a decrease in the absorption.

We think that, in the case of a rapidly decaying aperiodic discharge, e.g., in water, sparks of such a high intensity can be obtained that the detection of a decrease in the absorption in brightly fluorescent liquid solutions will also become possible.

Submitted in December 1925. The work was done at the Institute of Physics and Biophysics, Moscow.

14.3 Visual Measurements of Statistical Fluctuations of Photons: Translation of an Original Paper by E. Brumberg and S. Vavilov

Translated by Vladimir Rogovoi of the original paper by E. Brumberg and S. Vavilov (S. Wawilow): Visuelle Messungen der statistischen Photonenschwankungen (Visual measurements of statistical fluctuations of photons), *Izvestia Acad. Nauk SSSR, Ser. Math. (Bulletin de l'Académie des Sciences de l'URSS)*, OMEN, N 7, 919–941 (1933), in German. Reprinted in the book S. I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 345–364, in Russian.

Permission for publishing and translation from Nauka Publisher, Moscow, Russia

Submitted in July 1933. The work was done at the State Optical Institute, Leningrad.

14.3.1 Introduction

Until the present time, the corpuscular properties of light in the visible range of spectrum have been evidenced only indirectly, being theoretical inferences from numerous experiments. There were no direct observations of elementary statistical phenomena in this spectral range. Only in recently published paper by Barnes and Czerny [1], an attempt was made to detect fluctuations of photons using the dark adaptation of the eye. This attempt was based on the fact that near the threshold of visual perception of the human eye in the range of its maximal spectral sensitivity the flux of photons will be rarefied to such a high extent that such fluctuations should undoubtedly be expected.

The same idea underlay an unpublished work of one of us, which was conducted four years ago at the Moscow Institute of Physics and Biophysics and in which it was expected that the statistical variations of the setup with respect to the visibility threshold (in the König adaptometer) in the green light range would be larger than those in the red or violet range, since the threshold in the green range corresponds to a considerably smaller number of photons compared to the red and violet ranges. First tentative measurements, which were performed by V. I. Fedorova, yielded, in general, a positive result; however, the difference in the amplitudes of fluctuations proved to be small. In addition, there were grounds to worry about that the phenomenon is complicated by various psychophysiological factors. For reasons beyond our control, these experiments have been suspended for a long period of time, and only the recently issued above-mentioned paper by Barnes and Czerny prompted us to resume and continue our research in this field.

It seems to us that the observations by Barnes and Czerny were made under conditions in which it was extremely difficult to eliminate interfering physiological phenomena, which were pointed out by the authors themselves in their communication. A round surface, roughly 15 cm in diameter, with 50 randomly arranged holes was illuminated by green light and was observed from a distance of 50 cm without using the commonly applied red fixation point. The entire pattern was periodically visible only for a time period of 0.1 s, while during the remaining 0.9 s, it was screened shielded from the observer by a rotating sector. Therefore, most of the time, the extremely mobile and dark-adapted observing eye did not have any supporting point (the “fixation point” in physiological optics) and, clearly, could not permanently remain in the fixed position. In the absence of the fixation point, involuntary movements of the eyeball during the dark pause between flashes should inevitably lead to sharp variations in the observed brightness, in particular, because a rather large field of vision (about 18°) corresponded to different areas of the retina, the values of the scotopic sensitivity of which differed very strongly.

Another experiment by Barnes and Czerny, they described as follows: “One can use a lamp with a green filter that is applied in darkrooms for developing panchromatic plates. Initially, the light filter is covered by a thick layer of white paper and then it is covered with black paper that has appropriate holes. The brightness of holes can be controlled using a lamp rheostat with a wiper or by varying the distance to the head [of the observer]. Fluctuations can be seen only when the observer stays in the dark for a few minutes.” However, if we try to repeat this simple experiment, we can easily verify that “fluctuations” are quite noticeable not only under conditions of the visual threshold, but also at brightnesses that are hundreds of times higher than the generated ones. In other words, these observations have nothing to do with quantum fluctuations, and the phenomenon can be completely explained by physiological reasons.

Apart from that, we can state that, if all physiological effects were eliminated (except for one, namely: the property of the eye to retain the visual impression), all the same, quantum fluctuations in the experiment described above still cannot be observed. Due to the long duration of the visual impression, physical fluctuations should be smeared under the above conditions and be averaged in the same way as

they are averaged upon observation of threshold brightnesses from a large luminous surface. In this respect, the surface covered by isolated luminous spots does not have significant preferences over the homogeneously luminous surface of the same size. It should be noted that physiologists and psychologists have been familiar for a long time with the possibility of the human dark-adapted eye to observe fluctuations of luminous points; however, such fluctuations were always considered to be consequences of various physiological reasons [2].

It is difficult to understand this experiment by Barnes and Czerny, all the more so that, in their other experiments mentioned above, they applied the method of short light flashes.

From the energetic measurements by Barnes and Czerny, it follows that their final experiments were performed at such intensities that, theoretically, quantum fluctuations should be quite noticeable. However, their experimental conditions were such that physiological fluctuations should also definitely occur, with their magnitudes being larger, thus masking the presence of quantum fluctuations. Therefore, Barnes and Czerny, correctly indicating the possibility of visual observing quantum fluctuations, have not proved their occurrence in experiment either qualitatively or, much less, quantitatively.

The objective of this work is to discuss the question on the possibility to quantitatively observe quantum fluctuations. Based on previously obtained positive results, we perform experiments that can be considered to be the implementation of certain gedanken experiments of quantum physics. In the end of the work, we present certain inferences that are significantly important for physiological optics and photometry of weak light sources.

14.3.2 The Theory of Experiment

We assume that the threshold of sensitivity of the human eye exposed to light for a rather short period of time corresponds to a strictly determined number of photons n_0 . Experiments were performed provided that the following conditions are satisfied: (i) a quasi-monochromatic point light source is used; (ii) the position of the eye is permanent, which is achieved by applying a chinrest and a fixation point; and (iii) the eye is periodically excited for rather short time intervals with subsequent pauses of complete darkness (as in Barnes and Czerny's experiments).

Let n be the average number of monochromatic photons that were absorbed in the retina during the flash time and z be the actual number of photons that were perceived by the retina from this flash. By definition, only those flashes will be perceived for which

$$z \geq n_0. \quad (14.5)$$

In accordance with the results of different measurements of the threshold energy,

$$n_0 \gg 1. \quad (14.6)$$

The probability w that z lies between n_0 and ∞ provided that inequality (14.6) is satisfied, will be

$$w = \frac{1}{\sqrt{2\pi n}} \int_{n_0}^{\infty} \exp^{-\frac{(z-n)^2}{2n}} dz. \quad (14.7)$$

By substituting variable

$$y = \frac{z - n}{\sqrt{2n}} \quad (14.8)$$

we obtain

$$w = \frac{1}{\sqrt{\pi}} \int_{y_0}^{\infty} \exp^{-y^2} dy = \frac{1}{2} - \frac{1}{\sqrt{\pi}} \int_0^{y_0} \exp^{-y^2} dy = \frac{1}{2} - \frac{\varphi}{2}, \quad (14.9)$$

where φ is the Gauss integral, and

$$y_0 = \frac{n_0 - n}{\sqrt{2n}}. \quad (14.10)$$

Therefore,

$$\varphi = 1 - 2w. \quad (14.11)$$

The value of φ can be taken from tables. In our experiments, the value of the probability w was varied in the range from 0.15 to 0.8. In this interval, φ differs little from y_0 , as can be seen from Table 14.1.

For our purposes, everywhere where the high accuracy is not required, the following approximate value is quite sufficient:

$$w \sim \frac{1}{2} - \frac{y_0}{2} = \frac{1}{2} - \frac{1}{2} \frac{(n_0 - n)}{\sqrt{2n}}. \quad (14.12)$$

Let then

$$n = n_0 x, \quad (14.13)$$

Table 14.1

φ	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80
y_0	0.00	0.09	0.18	0.28	0.38	0.48	0.60	0.74	0.91

where x is the absorption coefficient, the value of which can be varied and controlled. From expressions (14.12) and (14.13) we have

$$w = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{n_0}{2}} \frac{(1-x)}{\sqrt{x}}. \quad (14.14)$$

The coefficient x can be smaller or greater than unity. In the latter case, the average number of photons n will be greater than the threshold number n_0 . At $x=1$, the probability w will achieve a value of 1/2, as this follows from expressions (14.11) and (14.14). This consequence is used for obtaining absolute values of x (cf. Sect. 14.3.4).

In our experiments, the maximum value of x was varied between 0.85 and 1.24, in accordance with which \sqrt{x} was varied in the range from 0.92 to 1.12. If it is necessary to further simplify results, it is acceptable to apply further simplification, which is obtained upon replacement of \sqrt{x} with unity. Then, expression (14.14) will acquire the form

$$w \sim \frac{1}{2} - \frac{1}{2} \sqrt{\frac{n_0}{2}} (1-x). \quad (14.15)$$

Therefore, in the interval of values of the variables w and x that we consider, there is a very simple relation between these quantities, which is expressed either by formula (14.14) or (14.15) and which can be used to quantitatively monitor the occurrence of quantum fluctuations. If the notion on the physical, quantum nature of flashes holds true, the probability w of perception of flashes within specified limits will linearly depend on $(1-x)/\sqrt{x}$ (or, roughly, on x), so that the slope of the straight line

$$\frac{d(w)}{d(\frac{1-x}{\sqrt{x}})} = \frac{1}{2} \sqrt{\frac{n_0}{2}} \quad (14.16)$$

will be a measure of the number of quanta corresponding to the threshold. At the same time, the passage to other wavelengths of the light source should be accompanied by changes in the slope of the curve such that they correspond to the spectral sensitivity of the eye (in the dark) and to the energy of light quanta. It can be easily verified that physiological fluctuations, which can be caused, e.g., by turns of the eye or by fluctuations in the refraction of its crystalline lens, can affect n and n_0 in a way that is complicated and poorly known to us, so that one cannot expect any simple (or of any kind whatsoever) relation between w , x , and the wavelength of the light source.

Finally, let us find the average intensity, i.e., the average number \bar{n} of photons in the flash at different values of x . The average value \bar{n} is expressed by the formula

$$\bar{n} = \frac{\frac{1}{\sqrt{2\pi n}} \int_{n_0}^{\infty} \exp^{-\frac{(z-n)^2}{2n}} dz}{w}.$$

After integrating and using formulas (14.8), (14.10), (14.11), and (14.14), we obtain

$$\bar{n} = \sqrt{\frac{2n}{\pi}} \exp^{-\frac{x^2}{2}} + n \approx \frac{1}{\sqrt{n}} \left\{ \sqrt{2n_0 x} + n_0(1-x) \right\} + n_0 x. \quad (14.17)$$

In the case of $x=1$, i.e., at $\varphi=0$, we have

$$\bar{n}_1 = \sqrt{\frac{2n_0}{\pi}} + n_0. \quad (14.18)$$

We know that, in this case, the average number of photons n is equal to the threshold number n_0 . In order to pictorially represent the dependence of \bar{n} on x , let us define the coefficient \bar{n}/\bar{n}_1 . Since $1-x$ did not exceed 0.2 in our experiments, we may replace $\sqrt{2n_0 x}$ in expression (14.17) by $\frac{1}{2}\sqrt{2n_0} + \frac{x}{2}\sqrt{2n_0}$. Then,

$$\frac{\bar{n}}{\bar{n}_1} \approx 1 + (x-1) \left\{ \frac{\left(1 - \frac{1}{\sqrt{\pi}}\right)}{\left(1 + \sqrt{\frac{2}{\pi n_0}}\right)} + \frac{1}{2(1 + \sqrt{\frac{\pi n_0}{2}})} \right\}. \quad (14.19)$$

Or approximately, provided that condition (14.6) is satisfied,

$$\frac{\bar{n}}{\bar{n}_1} \approx 1 + (x-1) \left(1 - \frac{1}{\sqrt{\pi}} \right) \approx 0.56 + 0.44x. \quad (14.20)$$

Since, as we already mentioned above, x varies in our experiments from 1.20 to 0.85, the average value of the energy of flashes should remain almost constant; however, at the same time, as is seen from (14.20), a considerable change in the number of flashes should be expected, as is seen from expression (14.16).

14.3.3 Last Experiments

Figure 14.10 shows the photograph of the setup of our last and final experiments. The setup was photographed from above. For individual components of the setup to be better distinguished in the photograph, both black tables on which the setup was arranged were covered with white paper upon shooting. As a light source, a four-volt incandescent lamp was used, which was fed from an accumulator. The lamp was placed into illuminator L_1 . The light passed through a milk glass, a green

filter, and a circular diaphragm with a diameter of 1 mm and entered tube R_1R_2 . To both ends of the tube, two polarization prisms were attached. By rotating these prisms with respect to each other, one could attenuate the light to a desired degree without changing its spectral composition. Between illuminator L_1 and tube R_1R_2 , sector disk S was installed, which was set into rotation with a synchronous motor with a transmission. The rotation rate of the disk was one revolution per second. The angular size of the slot in the disk was 36° . The center of the disk and the luminous point were arranged on the same vertical. The light from the second lamp L_2 was reflected by small mirror m , which was placed in front of the rotating sector disk. The image of a red luminous point lied in the same plane as the green point from the light source L_1 , at a distance of 60 cm from the eye. The observer saw the red point at an angular distance of 4° from the green one. The light from the red point was not interrupted by the rotation of the sector disk; hence, it remained always visible for the observer and was used as a fixation point. Between the sector disk and the tube R_1R_2 , a stack G of glass plates was placed, which was used to stepwise attenuate the light from the source L_1 . This step attenuator can be very easily used in the dark. Each glass plate attenuated the light flux by 7%, which was determined from photometric measurements.

On the second table, astronomic chronograph CB was placed with roll C of a paper chart and electrically controlled pens. These pens were connected with the rotating sector disk S in such a way that, to each revolution of the sector disk, a mark corresponded, which was automatically written with ink on the paper chart B . Using electric key K , which was placed on the first table, the observer could make his marks on the opposite edge of the moving chart.

Experiments were performed as follows. Preliminarily, the eye of the observer was subjected to the dark adaptation for the whole hour. To determine different stages of the adaptation, crystals of uranyl salts were used. Due to the radioactivity of uranium, uranyl salts are always excited by their intrinsic radioactive rays and emit weak light, which, however, can be rather clearly discerned by the dark-adapted eye. The intensities of this radiation are very different for different uranyl salts; therefore, one can easily form a scale from them with different steps of brightness. For example, a completely dark-adapted eye is capable of discerning the intrinsic light of uranium glass. These luminescence phenomena will be considered elsewhere.

The chin of the observer was immovably fixed by a special support (chinrest) A . In this case, the observer's eye could see through the polarization prisms R_1 and R_2 both luminous points L_1 and L_2 . The brightness of the points was controlled by the polarization prisms and, partially, by resistors W_1 and W_2 . The eye permanently fixes the red point, so that the green light is discerned only by the periphery.¹ In order for the eye position to remain unchanged in the course of observations, good training was required even in the presence of the fixation point. The state of the

¹Undoubtedly, at distances from the fixation point longer than ours, higher sensitivities could be achieved; however, experimentally, this is rather inconvenient.

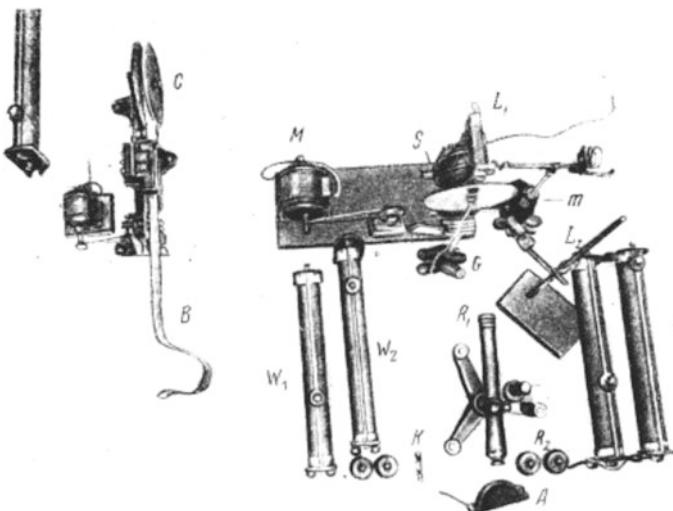


Fig. 14.10

observer also plays an important role (tiredness, before or after lunch observation time, etc.). Especially strongly, observations are interfered by the so-called intrinsic light of the retina, i.e., a visual sensation of light clouds that float in the dark. To a large extent, this phenomenon depends on the state of the observer; therefore, it is necessary to choose time periods for the observations such that this phenomenon either is least pronounced or is absent at all.

All final experiments were performed by one of us. In order to eliminate the effect of possible autosuggestion in the course of experiments, their results and calculations remained inaccessible to the observer for a long period of time.

Setting sector disk S into rotation, the observer fixed his eye such that to permanently see the red point and gradually reduced the brightness of the interruptedly emitting green point down to the threshold value (the fixation point was seen to be red, while the green point appeared to be colorless).

The first new qualitative result that we obtained is that, *in fact, there is no threshold of the visual sensation under these conditions!*

Initially, the observer sees a flash during each passage of the sector disk. Then, however, as the light is subsequently and very insignificantly attenuated, flashes cease to appear during each passage of the disk slot. During some passages of the slot, there appear no flashes, despite the fact that the brightness of each flash remains unchanged, at least, for the eye. If now the average light intensity of the source is gradually decreased further by inserting more glass plates into the stack G , flashes will become observed even more rarely, and, finally, they appear so rarely that the observer can see them only with a strong effort and, therefore, can easily miss them.

**Fig. 14.11**

For quantitative measurements, the observer should mark each flash (under conditions of the fluctuation regime) on the chart of the chronograph by pressing the key K . After a certain number of such marks (sufficient for statistical calculations), the observer introduces another glass plate into the stack G and marks again flashes that he sees on the chronograph chart. Performing such measurements with the use of four to five glass plates and provided that the dark adaptation was completely achieved requires one to two hours. Marks of flashes and of the number of passages of the disk slot that were made on the opposite edges of the paper chart yield all that is required to determine the probability w of the appearance of flashes. This probability is evidently equal to the ratio of the number of observed flashes (marks on one edge of the chart) to the number of passages of the slot of the sector disk (marks on the opposite edge of the chart). At the same time, the number of glass plates introduced into the path of the light beam serves as a measure of the relative value of x , i.e., the average number of photons corresponding to each flash.

In experiments with the red or violet light, the green filter in the illuminator L_1 was replaced by other filters. The transmission ranges of the red, green, and violet filters were 650–600, 520–490, and 460–425 m μ , respectively. We could not use a mercury lamp in our experiments due to inevitable variations in the intensity of this light source. Rapid oscillations in the intensity by a few percent could radically distort results.

14.3.4 Results

14.3.4.1 Statistical Randomness of Flashes

Each chart of the chronograph contains irregular marks of flashes and regular marks of periodic passages of the disk slot. Figure 14.11 shows a fragment of such a chart. The first test to which these protocols of observations should be subjected is to verify that marks of flashes are arranged randomly. If this condition is not fulfilled, this means that some systematic reasons affected the observation process.

Let w be the probability to observe a flash upon given passage of the sector disk slot. The probability p_n that, after the flash, there will be n passages without flashes and one passage with a flash is given by

$$p_n = (1 - w)^n w$$

or

$$\log p_n = n \log(1 - w) + \log w.$$

This formula can be used to monitor the statistical randomness of fluctuations. Let a_n be the number of intervals that correspond to n "dark" passages on the paper chart. Then,

$$p_n = \frac{a_n}{\sum a_n}.$$

Determining values of p_n by applying this formula to marks on the paper chart and taking their logarithms, we should obtain a straight line in the case of a statistical disorder. Processing all marks of flashes on the paper chart that we obtained from our observations yielded plots of straight lines. As an example, we present the data that refer to one of the charts (Table 14.2, Fig. 14.12).

The slope of the straight line in Fig. 14.12 characterizes the value of the probability w , which, in this example, is 0.52. At the same time, the value of the probability w can be found from the intersection of the straight line with the ordinate axis; in this case, $w = 0.50$. The directly calculated value of the probability is $w = 0.48$. Therefore, the marks on the chronograph chart are indeed distributed randomly.

Table 14.2

a	0	1	2	3	4	5	6	Number of passages 352
a_n	64	60	28	11	4	2	1	$\sum a_n = 170$
p_n	0.38	0.35	0.17	0.06	0.02	0.01	0.005	$w = \frac{171}{352} = 0.48$

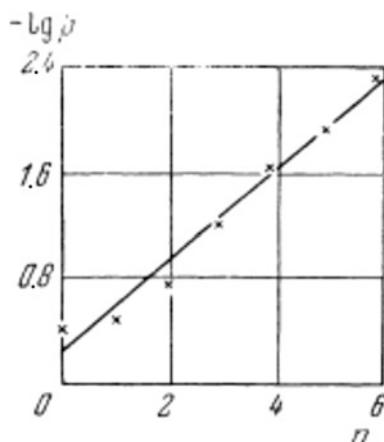


Fig. 14.12

14.3.4.2 Relationship Between w and x

In final experiments, we conducted five large series of observations altogether, in each of which the dependence of the number of observed flashes on the number of glass plates introduced into the light beam was determined. Results are presented in Tables 14.3, 14.4, 14.5, 14.6 and 14.7. In the first column of each table, G denotes the number of light-attenuating glass plates that were removed from the light beam. In the second column, α is the corresponding light transmittance expressed in relative units, with the initial value of this parameter being arbitrarily chosen to be unity. The third and fourth columns of the tables present values of x and $(1-x)/\sqrt{x}$, respectively. The method of determining x will be described further. In the fifth column, R designates the number of flashes, while, in the sixth column, N is the number of passages of the disk slot. Finally, the last column contains values of the probability

$$w = R/N.$$

All these five tables refer to the observations that were performed using the green light filter. The values of α that can be measured directly photometrically do not coincide with the values of x . Since, by formula (14.13),

Table 14.3

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	w
4	1.00	1.08	-0.077	48	69	0.69
3	0.93	1.01	-0.010	66	114	0.58
2	0.86	0.93	0.072	80	218	0.37
1	0.80	0.87	0.138	33	171	0.19

Table 14.4

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	w
6	1.00	1.24	-0.215	37	38	1.00
5	0.93	1.16	-0.148	26	32	0.81
4	0.86	1.08	-0.077	35	80	0.44
3	0.80	1.00	0.000	48	111	0.43
2	0.75	0.96	0.073	44	142	0.31
1	0.70	0.93	0.139	65	327	0.20

Table 14.5

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	w
4	1.00	1.11	-0.104	55	63	0.88
3	0.93	1.04	-0.039	70	125	0.56
2	0.86	0.96	0.041	61	15	0.39
1	0.80	0.89	0.117	57	236	0.24

Table 14.6

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	w
5	1.00	1.11	-0.104	79	100	0.79
4	0.93	1.04	-0.039	72	155	0.46
3	0.86	0.96	0.041	56	175	0.32
2	0.80	0.80	0.117	61	197	0.31
1	0.75	0.83	0.188	45	240	0.19

Table 14.7

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	w
4	1.00	1.04	-0.039	133	207	0.64
3	0.93	0.97	0.030	136	320	0.43
2	0.86	0.90	0.105	79	476	0.17
1	0.80	0.84	0.173	64	540	0.12

$$x = \frac{n}{n_0},$$

then

$$x = \frac{n}{n_0} = \frac{n'}{n_0} \alpha,$$

where n' is the average number of photons in the absence of light-attenuating glass plates. To determine the proportionality coefficient n'/n_0 , one can use the relation (cf. Sect. 14.3.2) in accordance with which, at $n=n_0$,

$$w = 1/2.$$

If this probability is reached, then

$$\frac{n'}{n_0} \alpha_0 = 1,$$

i.e., the transmittance coefficient α_0 , which corresponds to the probability $w=1/2$, is equal to the reciprocal value of the proportionality coefficient. To calculate values of x , we will proceed as follows. Initially, the plots of the dependence of the

probability w on the number of glass plates are constructed. Then, from an interpolation curve, which can always be presented in the form of a straight line, the value of the transmittance coefficient that corresponds to the probability $w=1/2$ is found. For Table 14.3, this operation is presented in Fig. 14.13. The first values in the columns for x in Tables 14.3, 14.4, 14.5, 14.6 and 14.7 are numerically equal to the proportionality coefficient n'/n_0 .

It is seen from these tables that this quantity always differs little from unity, which can be explained by the choice of the initial conditions of experiments.

The results presented in Tables 14.3, 14.4, 14.5, 14.6 and 14.7 are shown graphically in Figs. 14.14, 14.15, 14.16, 14.17 and 14.18. It is seen from them that experimental data, in accordance with theoretically calculated formula (14.14), yield clearly pronounced linear dependences of w on $(1-x)/\sqrt{x}$, despite the fact that, in some cases, there is a certain spread of experimental points. It should be noted that, at $(1-x)/\sqrt{x}=0$, all straight lines should pass through the point $w=1/2$, which restricts the arbitrariness of drawing of these lines. Large deviations of some points, as, e.g., those in Figs. 14.15 or Fig. 14.17, can be explained both by certain physiological reasons and by physical reasons, such as, e.g., unsteadiness of the source brightness. It is surprising that, in the majority of cases, points excellently fall on straight line. Based on the slope of straight lines, we can calculate by formula (14.16) the number of quanta n_0 , which corresponds to the sensitivity threshold of the observer in the blue-green range of the spectrum near 510 m μ . Table 14.8 presents values of the quantity

$$k = \frac{dw}{d\left(\frac{1-x}{\sqrt{x}}\right)} = \frac{1}{2} \sqrt{\frac{n_0}{2}}.$$

The values of this quantity correspond to Figs. 14.14, 14.15, 14.16, 14.17 and 14.18. It is seen from Table 14.8 that the threshold for the observer B, which was measured in different days over a period of a month, remained constant to an adequate degree. The absolute number of quanta n_0 fits well with Barnes and

Table 14.8

Figure number	k	n_0
5	2.4	46
6	2.2	40
7	2.7	56
8	2.2	40
9	2.6	54
Average	2.4	47

Czerny's measurements, who obtained a threshold of 40–90 quanta under conditions similar to ours (holes with a diameter of 0.4–2 mm, $\lambda \sim 530$ m μ , flash duration of 0.1 s).²

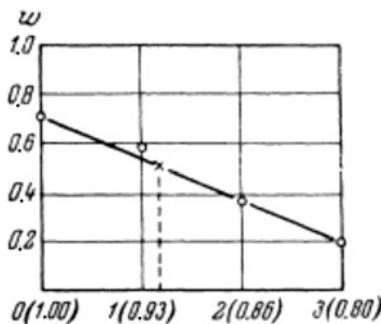


Fig. 14.13

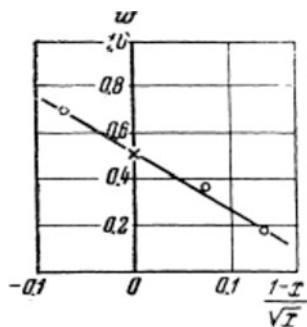


Fig. 14.14

²Based on our calculations of photons and results of energy measurements, the constant h can, in principle, be determined.

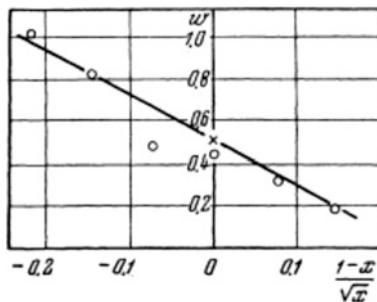


Fig. 14.15

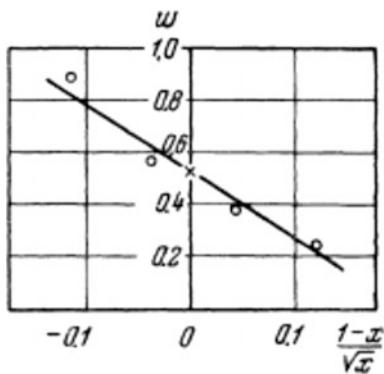


Fig. 14.16

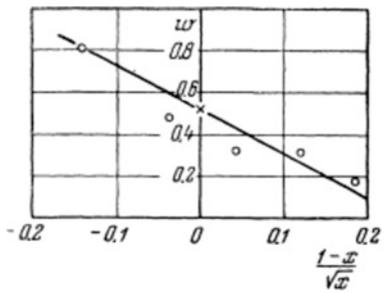


Fig. 14.17

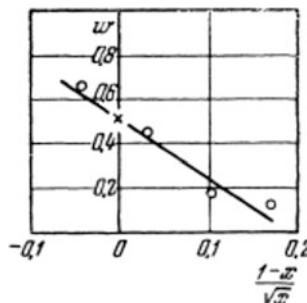


Fig. 14.18

14.3.4.3 The Sensitivity Threshold as a Function of the Wavelength

Because the sensitivity of the human eye falls rapidly in the long-wavelength and short-wavelength ranges of the spectrum, for quantitative measurements in these ranges, a monochromatic light source becomes necessary. Therefore, we could perform only qualitative experiments with our setup. These experiments confirmed completely our expectations. In experiments with the red filter, the intensity of the source was chosen such that, to almost every passage of the sector slot, a flash would correspond. In this case, it suffices to introduce only one glass plate for the probability w would drop almost to zero or, at least, to 0.01, which, theoretically, in accordance with formula (14.16), corresponds to a very large number of quanta n_0 . Consequently, in the red range of the spectrum in its critical range near the threshold, the eye proves to be sensitive to minimal changes in the light intensity. Similar results were also obtained with the violet filter, which can be seen from Table 14.9.

The number of observations performed was insufficient for drawing any definite quantitative conclusions from them, as this was done in experiments with the green filter. We can only state that the slope of the straight lines in these cases lies between 6 and 7, which corresponds to the number of photons n_0 roughly in the range 280–650. In connection with this, it should be noted that similar results could also be obtained in experiments with the green filter if the dark adaptation of the eye would be incomplete. This stationary state of the incomplete dark adaptation of the eye can be reached if the room is periodically illuminated for a short time with weak light. Such an experiment with the green filter showed that, upon introduction of only one glass plate into the light beam, the probability w dropped from unity almost to zero.

Table 14.9 Violet filter

G	α	R	N	w
3	1.00	32	38	0.84
2	0.93	14	64	0.22
1	0.86	5	86	0.06

All experiments described in this paragraph confirm satisfactorily the theory that was proposed in Sect. 14.3.2 and prove the possibility of visual observation of quantum fluctuations of the light flux.

14.3.5 Relative Fluctuations of Different Rays

Once the described method of direct visual measurement of quantum fluctuations was tested, it seemed interesting to us to apply this method to the investigation of coherent rays. It is well-known that the interference pattern can be photographed at extremely low intensities. This result is evident from the wave point of view and is completely unexplainable from the corpuscular point of view. Experiments that we will describe below show, however, that interference at low intensities is equally incomprehensible from the wave point of view as well.

For experiments, we used the same setup as before (Fig. 14.10), but with the difference that, between the rotating disk S and the tube R_1R_2 , a Fresnel biprism was placed, the refracting edge of which was directed horizontally. Therefore, in the field of vision, two coherent green points were seen, which were arranged symmetrically with respect to the red fixation point R (Fig. 14.19).

As soon as the threshold power was reached, the two green points fluctuated with respect to each other quite clearly and were very rarely seen simultaneously.

This phenomenon of *independent* relative oscillations of coherent rays has a catastrophic significance for the wave theory of light. This experiment admits only one explanation, namely, each ray interferes only with itself! If Fresnel could observe this such a simply reproducible phenomenon, clearly, the theory of light would be developed in absolutely another way!

Also, we studied this phenomenon quantitatively, determining the probability p of the simultaneous appearance of the two coherent points using the method described above. If the same average number of quanta corresponds to either of the two points, then, naturally, by the probability multiplication theorem, we have

$$p = w^2$$

or, using (14.14),

$$\sqrt{p} = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{n_0}{2}} \frac{(1-x)}{\sqrt{x}}. \quad (14.21)$$



Fig. 14.19

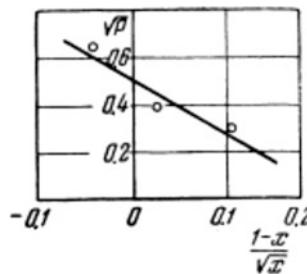


Fig. 14.20

Table 14.10 Green filter

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	p	\sqrt{x}
3	1.00	1.04	-0.039	40	107	0.37	0.61
2	0.93	0.97	0.030	35	201	0.17	0.41
1	0.86	0.90	0.105	14	155	0.09	0.30

Table 14.10 presents the results of one series of measurements (the notation is the same as in previous tables). Figure 14.20 shows these results graphically. In accordance with (14.21), we obtain a straight line, the slope of which is $k=2.3$. In accordance with Table 14.8, this corresponds to $n_0=43$ quanta.

A similar experiment was also performed with a Wollaston polarization prism, which was illuminated by a source of unpolarized (natural) light. In this case, two green points were polarized in two mutually perpendicular planes.

As in the preceding experiment, *the two points fluctuated absolutely independently from each other*. It follows from this observation that, at rather low intensities, the statistically natural light is polarized at each moment of time in a number of ways. It should be emphasized that this fact has nothing in common with “elementary polarization,” which is possible from the point of view of the wave theory and which should occur only within an extremely short moment of time. In our experiment, the state of a certain polarization of the natural light lasts for at least 0.1 s.

In this case, we also measured the probability p of a simultaneous flash of the two points as a function of x . Results of these experiments are given in Table 14.11

Table 14.11 Green filter

G	α	x	$\frac{1-x}{\sqrt{x}}$	R	N	p	\sqrt{x}
5	1.00	1.18	-0.165	56	66	0.85	0.92
4	0.93	1.01	-0.103	25	71	0.35	0.59
3	0.86	1.02	-0.02	30	100	0.33	0.57
2	0.80	0.95	+0.05	15	136	0.11	0.33
1	0.75	0.85	+0.16	4	108	0.036	0.19

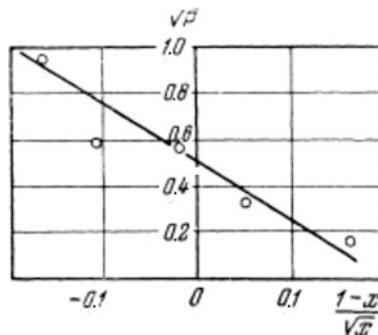


Fig. 14.21

and are shown graphically in Fig. 14.21. Here, we again have a rectilinear dependence of \sqrt{p} on $(1-x)/\sqrt{x}$, the slope of which is $k = 2.5$, which corresponds to 50 photons.

We could also realize various kinds of experiments with relative fluctuations of different rays. Thus, e.g., using a diffraction grating, we could easily observe relative fluctuations of the two first-order spectra. This would mean that the diffraction pattern does not arise immediately everywhere but, rather, emerges gradually, statistically. We could also implement experiments on observation of relative fluctuations of neighboring spectral lines of one and the same spectrum, which, statistically, would correspond to the distribution of energy over the spectrum. It is hardly worth dwelling on all possible variations of the experiments described above, because they all are based on one and the same common and empirically verified principle, which can be formulated as follows. *At a sufficiently low intensity, each ray exhibits statistical fluctuations, which occur completely intrinsically and independently of fluctuations in some other ray.*

14.3.6 Physiological and Photometric Consequences

Since the properties of light can be ascertained only via its action on some material detector, it is fundamentally impossible to determine with any experiment whether the exact reason for fluctuations described above lies in the light flux or in the detector. We can only state that all actions of the light are accompanied by these fluctuations, which, being quantitatively measured, exhibit a certain universal quantum character.

Apart from this physical inference, experiments described above point to some consequences, which may be of interest for physiological optics and, possibly, can be used as the starting point for the creation of new methods of research in the field of physiological optics.

If an interrupted point light source is used, as was the case in our experiments, then, strictly speaking, the sensitivity threshold will be the quantity that is unattainable experimentally. Upon lowering the intensity of the light flux, the brightness of flashes does not change appreciably at all; however, the number of flashes themselves changes significantly. Under these conditions, the sensitivity threshold n_0 cannot be determined directly energetically, at least with a sufficient accuracy. The only method to obtain n_0 can be provided only by statistical measurements similar to those that we applied. Therefore, on the one hand, one can hope to obtain the entire scotopic eye sensitivity curve, and with the advantage that the actual number of photons absorbed by the retina near the threshold of stimulation will be found. All energy losses that arise in the eyeball due to the reflection, absorption, and scattering, which inevitably affect any energetic methods, automatically fall away in statistical measurements. It is likely that, using this method, one can solve the very interesting question on the true sensitivity of the living retina in the long-wavelength ultraviolet range of the spectrum. As is known, ultraviolet rays in the range from 400 to 310 m μ are visible at high intensities, and the question remains open of whether the low sensitivity of the eye in this range is caused by the real insensitivity of the retina or by the absorption in the eyeball media.

Experiments with the red and violet rays that were presented in the end of Sect. 14.3.4 are of great interest for physiological optics, since they reveal an unexpectedly sharp boundary of the true sensitivity threshold n_0 .

In the majority of cases, the sensitivity thresholds are measured using surfaces that have large angular dimensions and that emit light for a long period of time; i.e., the measurements are performed under conditions under which statistical fluctuations prove to be smeared. In our experiments, the angular dimensions of the light source were still rather large (6'); however, it would be interesting to repeat these measurements using a light source the dimensions of which would correspond to only one element of the retina. The constancy of the sensitivity threshold of the observers eye during long-time experiments makes it possible to use the eye as an extremely sensitive instrument in the visible range of the spectrum for determining the relative intensity of very weak light sources, as well as for determining time changes of this intensity. In the case of a prolonged observation of a rather large surface (which can always be done by means of defocusing the optical system), when, consequently, statistical fluctuations are practically eliminated, this method can be applied in an especially simple form. To do this, it is only necessary to lower the brightness of the light source to the threshold value using any tool that is capable of measuring the degree of attenuation (e.g., a neutral wedge). Each change in the intensity of the light source will correspond to a certain change in the position of the wedge. Since the eye is a more sensitive detector in the green range of the spectrum than any other detector (photocell, photon counter), with the method that we proposed, one can measure the lowest intensities by the simplest means. A few years ago, one of us [3], has already applied this method, and, undoubtedly, it can be very successfully recommended for investigations of luminescence phenomena, in astronomical photometry, e.g., for the investigation of variable stars, and in the whole number of other fields. We will return to the detailed description of this

method in the future. Here, it is however worth emphasizing that the eye can be used, to some extent, as an absolute instrument for energy measurements.

Another remarkable property of the eye, which can be used for subtle measurements, is that the excitation function is virtually non-constant. As we saw (Sect. 14.3.4), in the range of critical intensities of the red and violet spectra, it suffices to change the transmittance coefficient by a few percent, for the probability w of flashes would drop from 1 to 0.01. Therefore, we can easily conclude that this phenomenon offers the opportunity to observe and measure minute changes in the absorption and intensity of light.

14.3.7 Conclusions

1. The constancy of the threshold of visual stimulation for the given observer during experiment and an extremely low energy magnitude of the threshold were applied to quantitative measurements of fluctuations of the light flux. After the critical introduction (Sect. 14.3.1), the theory of the method was given in the paper (Sect. 14.3.2).
2. Experiments totally confirmed the theory. Theoretical formula (14.14), which relates the probability w of light flashes to the extinction coefficient x via the threshold number of photons n_0 , was confirmed in all respects (Sects. 14.3.4 and 14.3.5).
3. The method was further used (Sect. 14.3.5) for the experimental detection of relative fluctuations of two coherent rays and polarized components of natural light. The theoretical suggestions were confirmed quantitatively.
4. Finally, some inferences that are important for physiological optics were considered and two new principles for absolute photometry of extremely weak intensities were proposed (Sect. 14.3.6).

References

1. R. Bowling Barnes and M. Czerny, Zs. f. Phys. **79**, 436 (1932).
2. C. E. Ferree, Amer. J. of Psychology **24**, 378 (1913).
3. S. I. Vavilov, Izv. Fiz. Inst. **1**, 96 (1920).

14.4 S. I. Vavilov's Comments to the Publication of S. Hecht, S. Shlaer and M. H. Pirenne: Translation from His Book "The Microstructure of Light"

This translation into English by Vladimir Rogovoi presents Vavilov's comments on the interpretation of his group's results by S. Hecht, S. Shlaer and M. H. Pirenne [Energy, Quanta and Vision, J. Gen. Physiol. **25**, 819 (1942)]. See

Chap. 17 (Sect. 17.3) with reprint of this paper. Hecht et al. carried out similar measurements to Vavilov's group on a sensitivity of a human eye to the faint light (see Chap. 17 of the current book). Vavilov's comment was published as a Footnote No. 2 in the original book in Russian S. I. Vavilov "The Microstructure of Light" (ML), Academy of Sciences of the USSR, Moscow (1950). See Chap. 2 (ML). Visual measurements of fluctuations abroad § 6 (ML). Measurements of fluctuations by S. Hecht et al.

Permission for publishing and translation from Nauka Publisher, Moscow, Russia.

Footnote No. 2 in the original book in Russian S. I. Vavilov "The Microstructure of Light" (ML), Academy of Sciences of the USSR, Moscow (1950)

The first communication by Hecht and his collaborators on fluctuation measurements has not contained references to our works published in 1933–1938 [1–5]. When publishing in 1942 our new fluctuation measurements, which were done in 1938–1941, we pointed out that our publications of previous years seemingly remained unknown to the American authors, since they were not been mentioned in their communication.

However, from a more detailed paper by Hecht et al., which was published in 1942 in the Journal of General Physiology, it became clear that our works were known to these authors. Moreover, these our works were briefly outlined by them in such a distorted form compared to the real content of our studies that we are forced to cite here completely the corresponding passage from their paper: "Barnes and Czerny (1932), and following them Brumberg and Vavilov (1933) realized that fluctuations must occur in the energy necessary for vision, and both groups of investigators looked for them. But they both missed the point of where the source of the fluctuations is and supposed it to be the energy deposited at the cornea. Brumberg and Vavilov even expected differences in the fluctuations for different wavelengths because of the greater energy required for seeing red light, for example, than blue-green light in conformity with the scotopic visibility curve."

In these lines, first of all, the statement of the American authors that our first work continues the study by Barnes and Czerny is surprising. It is clear from Sect. 2 (ML) that, from the experimental point of view, that work was erroneous, and only the theoretical indication on the necessity of the occurrence of quantum fluctuations was correct in it. The fundamental difference of our experimental work is evidenced, at least, by the circumstance that all main parts of our method and of our experimental setup [as it is clear from comparison of Sects. 3, 4, and 5 with Sect. 6 (ML)] were completely reproduced by Hecht and his collaborators. Our measurements are not a continuation of the work by Barnes and Czerny but rather they are a complete negation of the experimental part of their work.

Further, the American authors completely distorted the content of our work, alleging that we assumed that energy that is incident on the pupil of the eye fluctuates. On page 938 of our first work [1] (1933) [see translation of this paper in Sect. 14.3 of the current book (SGL)], they could read the following: [see Sect. 14.3.6 of the current book (SGL)] "The only method to obtain n_0 can be

provided only by statistical measurements similar to those that we applied. Therefore, on the one hand, one can hope to obtain the entire scotopic eye sensitivity curve, and with the advantage that the actual number of photons absorbed by the retina near the threshold of stimulation will be found. All energy losses that arise in the eyeball due to the reflection, absorption, and scattering, which inevitably affect any energetic methods, automatically fall away in statistical measurements. It is likely that, using this method, one can solve the very interesting question on the true sensitivity of the living retina in the long-wavelength ultraviolet range of the spectrum. As is known, ultraviolet rays in the range from 400 to 310 m μ are visible at high intensities, and the question remains open of whether the low sensitivity of the eye in this range is caused by the real insensitivity of the retina or by the absorption in the eyeball media." The lines presented above do not need any further clarification remarks. The American authors criticizing our work "overlooked" at best.

Finally, the American authors incorrectly ascribed to us the expectation of differences in fluctuations for different wavelengths. In the next paragraph (ML), it will be shown that this is not the expectation but rather the established fact, which was proven by measurements of many observers. For the unknown reason, Hecht and his collaborators restricted their consideration to the a priori statement that n_0 is independent of λ without presenting any measurement, despite the fact that there was a monochromator in their setup, which allowed them to easily pass to any range of the visible spectrum.

We have to protest against such an attitude toward works of other authors, which is extraordinary in regular scientific publications.

References

1. E. Brumberg and S. Vavilov (S. Wawilow): Visuelle Messungen der statistischen Photonenschwankungen (Visual measurements of statistical fluctuations of photons), *Izvestia Acad. Nauk SSSR, Ser. Math. (Bulletin de l'Académie des Sciences de l'URSS)*, OMEN, N 7, 919–941 (1933), in German. Reprinted in the book S.I. Vavilov, *Collected Papers, Vol. I. Papers on Physics of 1914–1936* (Academy of Science Publ., Moscow, 1954) pp. 345–364, in Russian. See Sect. **14.3** of this book with translation into English.
2. E.M. Brumberg and S.I. Vavilov (S. Wawilow): Statistical structure of interference field, *Dokl. Akad. Nauk SSSR (Soviet Physics: Doklady; Comptes Rendus de l' Académie des Sciences de l'URSS)* **3 (III)**, N 5, 322–325 (in Russian), 325–328 (in German) (1934).
3. S.I. Vavilov: Fluctuations of light and their measurements by statistical method, *Proceedings of the 1st Conference on Physiological Optics*, 25–29 December of 1934, Moscow-Leningrad, Academy of Sciences of USSR, 332–342 (1936), in Russian.
4. S.I. Vavilov: Directions of development of the Optical Institute, *Izvestia Acad. Nauk SSSR, (Bulletin de l'Académie des Sciences de l'URSS)*, Ser. Fiz. No. 1–2, 163–188 (1936), in Russian. See also the same paper in *Uspekhi Fizicheskikh Nauk (Physics-Uspekhi)* **16**, 872–896 (1936), in Russian.

5. S.I. Wawilow: Sensitivity of the retina to the ultra-violet spectrum, *Comptes Rendus (Doklady) de l' Académie des Sciences de l'URSS*, **21**, N 8, 373–375 (1938), in English and *Dokl. Acad. Nauk USSR (Soviet Physics: Doklady)*, 377–379 (1938), in Russian.

The editors suggest later papers of Vavilov and his group on sensitivity of the human eye to feeble light published either in English or in Russian (one of them is reproduced in Sect. 14.5)

1. E.M. Brumberg, S.I. Vavilov and Z.M. Sverdlov: Visual measurements of quantum fluctuations. I. The threshold of vision as compared with the results of fluctuation measurements, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), **VII**, N 1, 1–8 (1943), in English. The same paper in Russian: *ZhETF* **12**, N 3–4, 93–104 (1942). See Sect. 14.5 of this book.
2. S.I. Vavilov and T.V. Timofeeva: Visual measurements of quantum fluctuations. II. Fluctuations when the eye is light-adapted, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), **VII**, N 1, 9–11 (1943), in English. The same paper in Russian: *ZhETF* **12**, N 3–4, 105–108 (1942).
3. S.I. Vavilov and T.V. Timofeeva: Visual measurements of quantum fluctuations. III. The dependence of the visual fluctuations on the wave-length, *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), **VII**, 12–17 (1943), in English. The same paper in Russian: *ZhETF* **12**, N 3–4, 109–116 (1942).
4. S.I. Vavilov: Experimental investigation of quantum fluctuations of light by visual method, *Uspekhi Fizicheskikh Nauk* **36**, 247–283 (1948), in Russian.

14.5 Reprint of an Original Paper by E.M. Brumberg, S. I. Vavilov and Z. M. Sverdlov “Visual Measurements of Quantum Fluctuations. I. the Threshold of Vision as Compared with the Results of Fluctuation Measurements”

Reprinted with permission of Nauka Publisher (Moscow, Russia) from *J. Physics* (Publishing of the USSR Academy of Sciences, Moscow), **VII**, N 1, 1–8 (1943), in English. The same paper in Russian: *ZhETF* **12**, N 3–4, 93–104 (1942).

VISUAL MEASUREMENTS OF QUANTUM FLUCTUATIONS

I. The Threshold of Vision as Compared with the Results of Fluctuation Measurements

By E. M. BRUMBERG, S. I. VAVILOV and Z. M. SVERDLOV

(Received December 19, 1941)

On the basis of a method worked out earlier several series of fluctuation measurements were made with four observers for a wave-length $\lambda=525 \text{ m}\mu$ and an 8° angular distance of the fovea from the position of the retina under consideration. From a comparison with a black body the absolute value of the threshold of vision was determined from the same measurements. The average number of photons n_0 for each flash, corresponding to the threshold of the retina, was found to equal 25, and the average number of photons N , falling on the eye at each flash for the threshold conditions, was 208. The result agrees with the physiological data already obtained concerning the absorption of the eye media which confirms the quantum interpretation of the fluctuations. On the other hand, a comparison of n_0 and N opens up new possibilities for the physiological study of the properties of the media of the living eye.

In our former works (¹⁻³) experiments were described which show that it is possible to make visual observations and measurements of the fluctuations due to the quantum structure of light and the quantum character of absorption. The scope of the new measurements is a comparison of the number n_0 of photons, obtained for the threshold of the visual excitation of the eye from fluctuations, with the number N of photons, falling on the aperture of the pupil under the threshold conditions. The results obtained, at first glance unexpected, actually agree with the contemporary physiological knowledge of the optical properties of the media. Paragraph 1 contains the foundations of our method together with some remarks, corrections and additions, which were described in greater detail in an earlier work (¹). In the second paper of the present series are published the results of measurements of visual fluctuations with an eye adapted to a small light intensity instead of to darkness, and in the third—the results of measurements with different wave-lengths*.

* The present measurements made in 1938 and 1939 are a development of our former work which appeared in print in 1933 (¹⁻³). After having worked out our new results we read the short communication by S. Hecht, S. Schlaer and M. H. Pirenne (⁴) published in the June issue of the «Science», 1941. In this communication fluctuation measurements were described which agree

1. If the energy of the luminous flux is concentrated in photons distributed at random, then, for a sensitive receiver capable of responding to very weak luminous flux, fluctuations of perceivable energy must necessarily take place. As far as we know the only suitable receiver is a dark-adapted eye. An eye in this condition possesses a very great sensitivity and is capable of perceiving visually in the range near $510 \text{ m}\mu$ light intensities of the order of $10^{-10} \text{ ergs/sec.}$, i. e. about a hundred photons.

The observation of quantum fluctuations when the luminous flux is continuous is not possible because the visual sensation is very protracted and this causes an averaging of the fluctuations. It is impossible to make observations of fluctuations when the angular dimensions of the surface of the source of light are large. In conditions of peripheral vision and at very weak illumination the image on the retina is indistinct and the fluctuations at different places on a large image must mix up and on the average escape observation.

in many respects with ours in method, results and interpretation. Our publications were, evidently, unknown to the American authors as they did not mention them in their communication. In any case the agreement of the results in the given delicate experimental question is most satisfactory and is of essential importance.

Further it is necessary for the observations of the quantum fluctuations that the eye should be fixed, involuntary movements of the eyeball would inevitably cause sharp changes in the brightness having no connection with the physical fluctuations.

On the other hand, when the following three conditions are secured: 1) flash of short duration, 2) image of small dimensions on the retina and 3) fixation of the eyeball, the physical fluctuations must be observed if only our conceptions of the emission and absorption of light are correct and if the flux of light is sufficiently weak.

Qualitative and quantitative observations of the fluctuations are facilitated by the presence of a sharply defined threshold of visual excitation. In the absence of such a threshold the fluctuations would result in fluctuations of the brightness of the observed flashes. Owing to the existence of a threshold those flashes which have an energy less than the threshold value are simply invisible. Consequently, there is a very sharp indication of the existence of fluctuations, the flashes being either visible or not; an insignificant change in the brightness of the flashes is practically unnoticeable.

Let n denote the average number of monochromatic photons absorbed by the retina for one flash of short duration, and let z be the actual number of photons absorbed during the given flash. Owing to the presence of a threshold only flashes for which

$$z \geq n_0 \quad (1)$$

will be perceptible.

Below it will be assumed that in the given experimental conditions n_0 has a strictly constant value. If in the actual conditions n_0 really oscillates irregularly from flash to flash about a certain average value (for physiological reasons) then in the final formulae (see below*) n_0 is equivalent to this average value. More or less marked variations of n_0 are also possible from one group of obser-

* It is possible that n_0 is determined not by the maximum number of absorbed photons, but by the number of rods having absorbed at least one photon and therefore taking further part in nervous excitation. From this point of view a rod which has absorbed two or more photons does not, in respect of its action, differ from a rod having absorbed one photon. If this is so, the absorption of n_0 photons is not yet equivalent to the threshold excitation, and therefore changeable n_0 is in reality a certain mean value.

vations to another owing to various physiological reasons. The variations of n_0 must also give rise to a divergence between the observed probabilities and the theoretical ones which is one of the reasons for the dispersion of the experimental points.

From the theoretical point of view the character of the fluctuations of the light flux is not up to the present time quite clear even in the case of radiative equilibrium (*). In the case of a directed light beam such as those used in the experiments under consideration, the lack of certainty is even greater. Besides this, we are, in fact, inevitably concerned only with the fluctuations of absorbed light energy. It is only under conditions of total absorption, when the number of incident photons and those absorbed in the medium is the same, that the fluctuations in the light flux and in the absorbing medium must coincide. When the absorption is not total the character of the fluctuations must, generally speaking, depend on the fluctuations in the incident light beam. It is clear, for example, that if the incident beam is regular and there are no fluctuations in it at all, the statistics of absorption will differ from that in cases of irregular fluctuations in the incident beam. In connection with this lack of clarity in the theory of fluctuations an experimental study of the statistical processes when light is absorbed is particularly important.

We suggest that it is possible to determine the statistical processes under consideration by means of Poisson's formula. The successive modulation of the one fluctuations by the other ones in this case cannot change the Poisson character of the fluctuations.

By supposing that, according to experiment:

$$n_0 \gg 1 \quad (2)$$

for the probability of a certain value z when n is the average value according to Poisson's formula:

$$p(z) = \frac{n^z}{e^n z!} \quad (3)$$

Under the condition (2) the probability p that z has some value between n_0 and ∞ can be expressed by the integral:

$$p = \frac{1}{\sqrt{2\pi n}} \int_{n_0}^{\infty} e^{-\frac{(z-n)^2}{2n}} dz. \quad (4)$$

VISUAL MEASUREMENTS OF QUANTUM FLUCTUATIONS. I

By introducing

$$y = \frac{z-n}{\sqrt{2n}} \quad (5)$$

(4) can be rewritten as

$$\begin{aligned} p &= \frac{1}{\pi} \int_{y_0}^{\infty} e^{-y^2} dy = \\ &= \frac{1}{2} - \frac{1}{1/\pi} \int_0^{y_0} e^{-y^2} dy = \frac{1}{2} - \frac{\varphi}{2}, \end{aligned} \quad (6)$$

where φ denotes the probability integral, and

$$y_0 = \frac{n_0 - n}{\sqrt{2n}}. \quad (7)$$

For the probability interval 0.4, ..., 0.9, which is of predominance in experiments, φ does not differ within the limits of experimental error from y_0 and (6) can be approximately written as

$$p \approx \frac{1}{2} - \frac{y_0}{2} = \frac{1}{2} - \frac{1}{2} \frac{(n_0 - n)}{\sqrt{2n}}. \quad (8)$$

If n_0 varies itself statistically from flash to flash, owing to physiological reasons, then, from formula (8) due to its linear character with respect to n_0 , it follows that n_0 in the expression for p must be substituted by the average value \bar{n}_0 . In other words, fast statistical variations of n_0 will not change the results and qualitative conclusions made up to now.

Let us denote

$$\frac{n_0}{n} = x, \quad (9)$$

then

$$p = \frac{1}{2} - \frac{1}{2} \sqrt{\frac{n_0}{2}} \frac{(1-x)}{\sqrt{x}}. \quad (10)$$

It is convenient to determine the absolute value of x experimentally from the condition:

$$\text{for } x = 1 \quad p = \frac{1}{2}, \quad (11)$$

which follows from (6).

By finding experimentally the values of p and the relative values of x , we can determine the absolute values of x graphically from the condition (11) after which a curve $p=f(x)$ can be plotted.

On the basis of the more exact formula (6) the curve should look like that represented in Fig. 1 where p and $\frac{1-x}{\sqrt{x}}$ (which is pro-

portional to y_0) are chosen as axes. Formula (10) is valid for the central rectilinear part of the curve in Fig. 1. The slope k of the rectilinear part is

$$k = \frac{1}{2} \sqrt{\frac{n_0}{2}}, \quad (12)$$

wherefrom

$$n_0 = 8k^2. \quad (13)$$

The main object of the first part of the present paper, as was mentioned, consists in

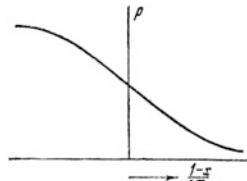


Fig. 1

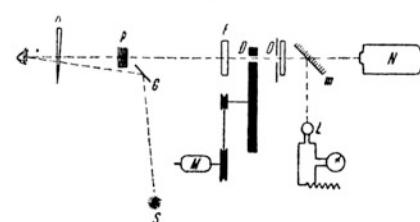


Fig. 2

a comparison of the number of photons n_0 determined from the fluctuations with the number of photons N corresponding to the threshold of vision, this number being determined from the energy falling on the pupil of the eye under the threshold conditions.

2. The scheme of the apparatus by means of which the measurements were carried out is shown in Fig. 2. The head of the observer rests on his chin. The eye is fixed on the red point S (the brightness of which may be decreased at will by means of the rheostat) by means of the glass plate G . By varying the slope of the plate and its distance from the axis of the apparatus one can vary the angle ω between the fovea and the tested place of the retina. To this place comes the light from the

source passing through diaphragm O covered with a milk glass, through the opening in the disk D , brought into rotation by means of a synchronous motors M , the light filter F , the pile of glass plates P and the photographic wedge K . The diaphragm O is illuminated by the incandescent lamp L fed by an accumulator, and the light is directed into the eye by the mirror m . For absolute measurements of energy the mirror and the milk plate in front of the diaphragm O are taken away and the diaphragm is illuminated by the black body N .

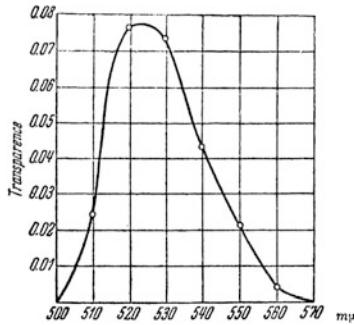


Fig. 3

Before starting observation the eye is adapted to total darkness for one hour, after which the motor is started and the disk makes one revolution every 1.6 sec. The indentation in the disk was 20° , i. e. the duration of the flash was $9 \cdot 10^{-2}$ sec. The brightness of the light was lessened by the wedge until it was observed that flashes began to remain unnoticed at the passing of the indentation in the disk. For a very bright light there is a visible flash corresponding to each passage of the opening. Omission are interpreted as the result of fluctuations in the luminous flux and in the absorption. A fine regulation of the brightness, i. e. of the value of x in formula (10) is made by the observer in darkness by switching the glass plates in the pile on and off. Each plate diminishes the intensity of light by approximately 7%. The method of the pile makes a qualitative variation of x possible without any calculating and registering in the darkness.

The key of the chronograph tape is placed by the hand of the observer. On one edge of,

the tape each revolution of the disk is automatically registered and on the other edge are registered (when pressing the key) the observed flashes. The fraction obtained by dividing the number of observed flashes by their general number gives the probability p . However, if during the experiment the number of flashes is small, the calculated probability may be very different from the a priori probability and this circumstance must influence the distribution of the experimental points.

The area of the round opening in the diaphragm O was $3 \cdot 1 \cdot 10^{-2} \text{ cm}^2$ and its distance from the eye was 116 cm, so that the angular distance of the observed lightspot was $6'$. The observations were made at a distance of $8'$ from the fovea.

The filtration of the filter F is given in Fig. 3. The transmitted range was sufficiently narrow, what could be qualitatively verified by the observations of the colour of the transmitted rays when the light from the lamp was continuously decreased. The light remained green the whole time with no noticeable change in the hue. The transparency curve has its centre of gravity at about 525 m μ .

By means of this apparatus and on the basis of the method described in paragraph 1, n_0 can be determined, i. e. the minimum number of photons which, on being absorbed by the retina, cause a visible sensation. To find N , i. e. the number of photons incident on the pupil when n_0 photons are absorbed by the retina, the energy incident on the pupil during the flash must be known. To this purpose the black body N (Fig. 2) was used, which could be moved up to the diaphragm (when the milk glass has been taken away). By knowing the temperature θ of the black body, the dimensions of the diaphragm O , the spectral transparency of the filter F , the distance of the diaphragm O from the pupil and the diameter of the pupil of the observer under the conditions of darkness adaptation, one can determine, by using the laws of the black body radiation, the intensity of the luminous flux, incident on the pupil when n_0 photons are absorbed by the retina.

The temperature of the black body was about 1100°K . It was measured by a platinum-iridium thermoelement, graduated in All-Union Institute for Metrology and Standards with an accuracy of 2° and besides this controlled by comparing it with the total radiation of a Hefner lamp. The differences in the temperatu-

VISUAL MEASUREMENTS OF QUANTUM FLUCTUATIONS. I

res determined by the thermoelement and the Hefner lamp did not exceed 3–6° when the measurements were made as exactly as possible and were of an irregular character. For rougher measurements the difference reached once as much as 40°. The reason for such a large difference was probably connected with the habitual inconstancy of measurements with a Hefner lamp and in any case

The diameter of the fully adapted pupil was determined by photographing the eye by the light from a flash-lamp. A scale was pressed to the eye by means of which the absolute size of the adapted pupil was found.

As it was inconvenient to work continually with the black body, the lamp L in a definite regime (accumulator control apparatus and light filter F) was compared with the

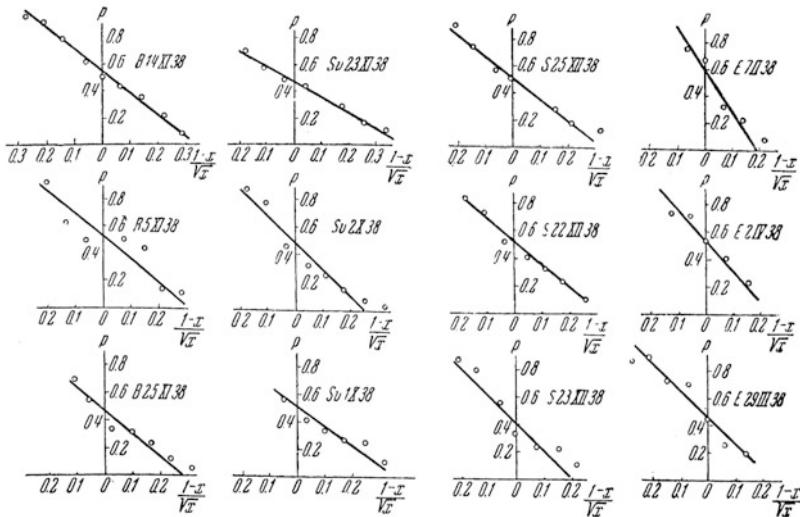


Fig. 4

must, most probably, be put down to the Hefner lamp rather than to the thermoelement.

For that spectrum and temperature range where Wien's law of energy distribution holds, the connection between a small change ΔE of the energy of the radiation and a small change of the temperature $\Delta\theta$ is expressed by the formula

$$\frac{\Delta E}{E} = \frac{hc}{\lambda k_0} \frac{\Delta\theta}{\theta}$$

Let us suppose that $\Delta\theta$ is 5°, then $\lambda = 5.25 \cdot 10^{-5}$ and $\theta = 1.1 \cdot 10^3$ equals approximately 10%. In other words, the accuracy of the value of the energy N given below is about 10%.

black body. Further on it was used for relative and absolute measurements.

According to formula (41), when $p = 1/n$, $n = n_0$. If, besides the curve $p(x)$ we know the absolute value of the energy for a certain x , obtained, as was already pointed out, by a comparison with a black body, then the absolute value of x can be found and in this way N is determined from x , corresponding to $p = 1/n_0$. It is important to notice that in our experiments N was found in this way from the same measurements and under the same conditions as n_0 .

3. Four observers took part in the measurements: B , Sv , S , E , each fulfilling 5–10 series of observations. Fig. 4 gives examples of typical series for all the observers. On the

axis of the abscissae the value of $(1-x)/\sqrt{x}$ is given and along the axis of the ordinates the probability p . From this drawing we can form an idea of the way in which the experimental points are scattered. It is necessary to point out that it is only possible to obtain good results after a long preliminary training of the observers which is only obtained after 5–10 sittings lasting at least one hour each. The aim of the training is to habituate the eye to remain fixed, to give it a training in the peripheral vision and also to habituate it to some concentration of attention which is necessary for registering all the observed flashes by pressing the key.

The first measurements of all observers (before training) have a disorderly character showing practically none of that linear regularity which clearly makes its appearance later, after training.

The method of working out the observations is described in the first part of this paper⁽¹⁾.

As in all our former observations, the large amount of material in the present work, including about 50 series (about 50 000 separate calculations), fully confirms the presence of an approximately linear dependence between p and $(1-x)/\sqrt{x}$ which follows from our theory.

Table 1 gives the main results of the present work. For each observer are shown the data of observations, the number of photons N , falling on the average on the eye during each flash, when $p = \frac{1}{2}$, and the slope of the curve k obtained by a graphic interpolation of the measurements.

By using formula (13) we find from Table 1 the comparative values of N and n_0 (Table 2).

The results of the fluctuation measurements, collected in Tables 1 and 2 lead to the following conclusions:

1) For the given experimental conditions ($\lambda = 525 \text{ m}\mu$, $\omega = 8^\circ$) n_0 oscillates for the four observers within the limits 18–32 and equals on the average 25 photons. From the observations of one of us (B) made in 1933 we find that $n_0 = 47$ photons⁽¹⁾, the experiments being carried out with a green and wider filter. The monochromatic experiments of 1936 when $\omega = 10^\circ$ for $\lambda = 510 \text{ m}\mu$ gave, for the same observer B , $n_0 = 30$, and for another P , $n_0 = 8^*$.

* The American observations cited above for $\omega = 20^\circ$ give on an average $n_0 = 6$ for three observers.

Table 1

Date	N	k	Date	N	k
Observer B			Observer Sv		
2. X. 1938	154	1.7	2. X. 1938	333	1.4
4. X	150	1.5	5. X	324	1.8
5. XI	173	1.7	10. X	323	1.4
14. XI	178	1.5	22. X	335	1.4
16. XI	145	2.3	20. XI	304	1.7
25. XI	178	1.5	23. XI	318	1.4
28. XI	179	2.0	27. XI	333	1.5
29. XII	191	1.6			
Mean value	168	1.7	Mean value	324	1.5
Observer S			Observer E		
16. XII. 1938	224	2.0	4. II. 1938	—	2.1
20. XII	192	1.8	5. II	—	1.9
22. XII	291	1.8	7. II	—	2.1
23. XII	255	1.9	29. III	113	1.8
25. XII	246	1.5	2. IV	84	1.9
26. XII	228	1.9	3. IV	103	2.0
28. XII	242	2.3	4. IV	131	1.9
29. XII	221	1.6			
30. XII	178	2.2			
Mean value	231	1.9	Mean value	108	2.0

Table 2

Observer	B	Sv	S	E	Mean value
Mean value of N	168	324	231	108	208
Mean value of n_0	23	18	29	32	25
$\frac{N_{av}}{n_{0av}}$	7.3	18	8.0	3.4	8.3

So we see that the oscillations of the values of n_0 for different observers can be very large.

According to our observations we find that for the same observer on different days n_0 oscillates (in rare cases) about two times. However, as the slope of the straight line k cannot be determined very exactly (for reasons given in paragraph 1), there is not sufficient basis to consider the differences in the values of n_0 given by different observers or by any one observer on different days as real ones. This is the more true as the value k to which the process of interpolations is applied has to be squared in order to obtain n_0 .

2) The value of N oscillates for the same observer on different days much less than that of n_0 , as can be seen from Table 1. At the same time for different observers the values of N differ very considerably, three times. Such a divergence cannot be ascribed to random causes.

3) From Table 2 it follows that $N \gg n_0$, on an average $N = 8.3 n_0$; for different observers the ratio N/n_0 shows very large variations, more than five times.

rely cover the surface of the retina especially in regions far from the fovea. Fig. 5 gives the curve showing the distribution of the rods at different distances from the fovea, according to Osterberg's data (*). According to Osterberg, the usual number of rods on the retina is about $1.2 \cdot 10^6$ and the diameter of the rods is about 2μ . The rods, according to Osterberg, are distributed over an area of about 10 cm^2 , i. e., on an average the rods take up rather more than one

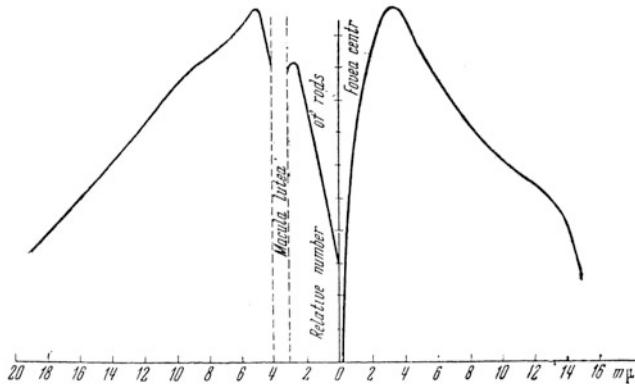


Fig. 5

4. The very significant difference between n_0 and N which is the main result obtained appears at a first glance to be in strong contradiction to our principal suggestion concerning the quantum character of the observed fluctuations. Such a conclusion, however, is not correct. On the contrary, modern data concerning the optical properties of the eye medium necessarily demand that N should be much larger than n_0 .

Laying aside the small loss of light owing to the reflection from the eye surface it is necessary to take into account the absorption of light in the eye medium before coming to the retina. According to the data of Roggenbau and Wetthauer (*), the absorption of light in the region of about $510 \text{ m}\mu$ equals for all the eye media about 50% before reaching the retina. On the other hand, the rods which absorb the light giving rise to the sensation of vision do not enti-

third of the whole surface of the retina. In the region corresponding about $1.6 \cdot 10^6$ rods are distributed over the eye retina, consequently, only 0.5 of the area at this point is covered by the rods.

Further on the surface of the retina itself we find nerves and other tissues with a certain unknown absorption of the incident light.

Finally, the absorption of the rods themselves in the purple vision even when at its maximum is not total. According to Wald's (*) data, the coefficient of absorption of the visible purple is about $500 \text{ m}\mu$ per unit of the surface of the eye retina: for frogs it is 33.6%, for rabbits—4.2% and for rats—13.0%.*

* Wald points to the fact that his figures are definitely lower owing to the incomplete extraction of rodopsine. In the subsequent Wald purely arbitrarily assumes the absorption coefficient equal to 25%.

By α let us denote the transmission coefficient of light in the eye media before attaining the retina; by β —the transmission coefficient in those layers of the retina itself, by which the rods are screened; by γ —the portion of the area of the retina covered by the rods in the region under consideration, and by δ —the absorption coefficient of light in the rods. Then it is obvious that

$$n_0 \geq N \alpha \beta \gamma \delta. \quad (14)$$

The sign $<$ corresponds to the possibility that the output of the physico-chemical processes determining the act of vision is < 1 .

By using the figures given above for the conditions of our experiments we can at any rate write

$$n_0 \geq N \frac{1}{2} \cdot \frac{1}{2} \beta \delta,$$

i.e.

$$\frac{N}{n_0} \geq \frac{4}{\beta \delta}.$$

By comparing the average experimental results with $N/n_0 = 8.3$, we come to the conclusion that

$$\beta \delta \approx \frac{1}{2}.$$

As has already been mentioned there is nothing known about the value of β and in any case it is probable that β differs little from unity; as far as δ is concerned, according to the estimation of the physiologists, its value in the region of about 500 m μ is not less than 0.2. This shows that our results are in qualitative agreement in so far as all the quantitative knowledge of the eye media is very indefinite and arbitrary and has only the character of estimations.

It is clear that all the four parameters α , β , γ , δ can vary within wide limits for

different people, whereas at the same time their considerable variation in a short interval of time for one and the same person is not very probable. This conclusion agrees with experiment as can be seen from Tables and 2.

In this way, besides the confirmation of the principal hypothesis of the quantum character of the observed fluctuations, we get a result which is of essential interest to the study of the eye. The simultaneous determination of n_0 and N by the proposed method (or a variation of it) makes it possible to measure for the living eye separately the sensitivity of the retina and the share of light actively participating in the act of vision.

However this fraction is given by measurements of n_0 and N only in the undifferentiated form as the product of $\alpha \beta \gamma \delta$. Making an additional supposition regarding the properties of the eye and visual perception it would seem possible to reach certain conclusions at least with reference to γ .

It is well known that the sensitivity of peripheral vision depends upon ω , increasing several times on passing from 0 to 10° (according to the most dependable data) and then falling down. A plausible explanation of this dependence consists in a corresponding change in the concentration of rods per unit area (cf. Fig. 5). If such an explanation is correct, change of N as depending upon ω is fully determined by variation of γ , while, on the other hand, n_0 should not depend on ω . No systematic experiments with different ω for one and the same observer have so far been carried out. It is hardly possible to bring together the experiments of different observers on account of the very considerable differences of n_0 for different observers.

State Optical Institute.

REFERENCES

- ¹ E. Brumberg u. S. Wavilow, Bull. Acad. Sci. URSS, série math., p. 919 (1933).
- ² E. M. Brumberg a. S. I. Wavilow, C. R. Acad. Sci. URSS, **3**, 1 (1934).
- ³ С. И. Вавилов, Труды конференции по физиологии оптики (С. И. Вавилов, Trans. of the Conf. on Physiol. Optics), Leningrad, 1936.
- ⁴ S. I. Wavilow, Bull. Acad. Sci. URSS, série phys., No. 1—2, 176 (1936).
- ⁵ S. I. Wavilow, C. R. Acad. Sci. URSS, **21**, 377 (1938).
- ⁶ S. Hecht, S. Schlaer a. M. H. Pirenne, Science, June 20, 585 (1941).
- ⁷ A. Einstein, Phys. ZS., **10**, 185, 817 (1909); W. Heisenberg, Ber. d. Sächsisch. Akad. Wiss., **88** (1931); M. Born a. K. Fuchs, Proc. Roy. Soc., A, **170**, 252 (1939); **172**, 465 (1939).
- ⁸ C. Roggenbau u. A. Wetthauer, Klinisches Monatsblatt der Augenheilkunde, **78**, 762 (1927); **79**, 456 (1927).
- ⁹ G. Osterberg, Acta ophtalmologica, Supplementum VI, Copenhagen, 1935.
- ¹⁰ G. Wald, Journ. of Gener. Physiology, **22**, 765 (1938).

Chapter 15

Nonlinear Optical Experiment of 1941 (Gilbert Newton Lewis)



Svetlana G. Lukishova

Abstract This chapter is devoted to a prominent chemist Gilbert Newton Lewis (1875–1946) and a nonlinear optical experiment (saturation of absorption and phosphorescence) performed by him and his group (D. Lipkin and T. T. Magel) in 1941. Section 15.1 contains a brief description of Lewis' biography, and Sect. 15.2 outlines his paper of 1941 on saturation of absorption and phosphorescence. It should be mentioned another Lewis' contribution to optics: after his usage of the word “photon”, this name for “light quanta” became very popular in scientific literature. Lewis is best known for his work on chemical bonding and the production of the first sample of “heavy water”.

Gilbert Newton Lewis, one of the most influential scientists of the last century, made significant contributions in both chemistry and physics [1–2]. He was nominated 41 times for the Nobel Prize in chemistry [3] and was the mentor of 20 Nobel Prize laureates and 290 Ph.D. recipients [2]. The main directions of his research and achievements were concentrated on chemical thermodynamics, valence theory (the concept of the covalent bond and “Lewis symbols”), the electronic theory of acid and bases, production of the first sample of “heavy water” (deuterium oxide), etc. Although he was not the first to introduce the term “photon” (see preface of this book), it was after his use of the word in his Nature paper, that it became very popular in a scientific literature.

In this chapter we describe Lewis' pioneering experiment on nonlinear optics (1941) on saturation in absorption and phosphorescence, that was carried out before the laser invention, using a high-pressure, 1000 W mercury arc. He started to work on dye photochemistry/photoluminescence, in particular on the triplet state and phosphorescence, from the late 1930s which he continued until his death in 1946.

This chapter contains two sections. Section 15.1 is a Gilbert Newton Lewis' biography reprinted partially with permission from American Physical Society

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road,
Wilmot Building, Rochester 14627, USA

e-mail: sluk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

News from <https://www.aps.org/publications/apsnews/201212/physicshistory.cfm>. Section 15.2 is the editor's overview of an original publication on nonlinear optical effects in phosphorescence and absorption by G. N. Lewis, D. Lipkin and T. T. Magel: Reversible photochemical processes in rigid media. A study of the phosphorescent state, *J. Am. Chem. Soc.* **63**, 3005–3018 (1941). A text excerpt and figures from the original paper are included in this overview with permission of the Americal Chemical Society. In this work G. N. Lewis et al. used fluorescein in boric acid as a highly nonlinear medium with a phosphorescence decay time up to a few seconds.

We also like to mention Chap. 14, Sect. 14.2 of this book, with translation into English of the excerpt of the earlier paper (1926) of Vavilov and Levshin on nonlinear optics (saturation of absorption using uranium glass). Chapter 16 of this book describes a nonlinear optical device by the Institute of Optics, University of Rochester, based on saturation of fluorescence and employed by US pilots during Word War II.

Although Lewis et al. used intense light from a high-pressure mercury arc (~ 1000 W), in a recent publication on the same material (a fluorescein-doped boric-acid glass) [4] saturation of absorption was reported at very low intensity of ~ 15 mW/cm², and a nonlinear susceptibility $\chi^{(3)}$ was evaluated as large as ~ 1 esu, as compared to $\sim 10^{-12}$ esu for the commonly used Kerr liquid CS₂.

- [1] J.H. Hildebrand: *Gilbert Newton Lewis*, 1875–1946, Biographical Memoir, National Academy of Sciences, Washington, DC (1958). <http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/lewis-gilbert-n.pdf>
- [2] Historical Inventors: Gilbert Newton Lewis <https://lemelson.mit.edu/resources/gilbert-newton-lewis>
- [3] https://www.nobelprize.org/nomination/archive/show_people.php?id=5441
- [4] M.A. Kramer, W.R. Tompkin, and R.W. Boyd: Nonlinear-optical interactions in fluorescein-doped boric acid glass, *Phys. Rev.* **34**, 2026–2031 (1986).

15.1 Gilbert Newton Lewis: Biography

Reprinted partially (with permission from American Physical Society News) (<https://www.aps.org/publications/apsnews/201212/physicshistory.cfm>).

A native of Weymouth, Massachusetts, Lewis (Fig. 15.1) was largely educated at home by tutors, although he briefly attended public schools between the ages of 9 and 14, at which point he enrolled at the University of Nebraska. Three years later, at age 17, he transferred to Harvard College, completing his PhD in 1899 when he was just 24. He spent a year in Germany as a traveling fellow, studying under notable scientists like Wilhelm Ostwald and Walther Nernst, before returning to Harvard to teach.

In 1904, Lewis moved to the Philippines to head the Bureau of Science's unit on weights and measures, but he held the post for only a year. Instead, he returned to

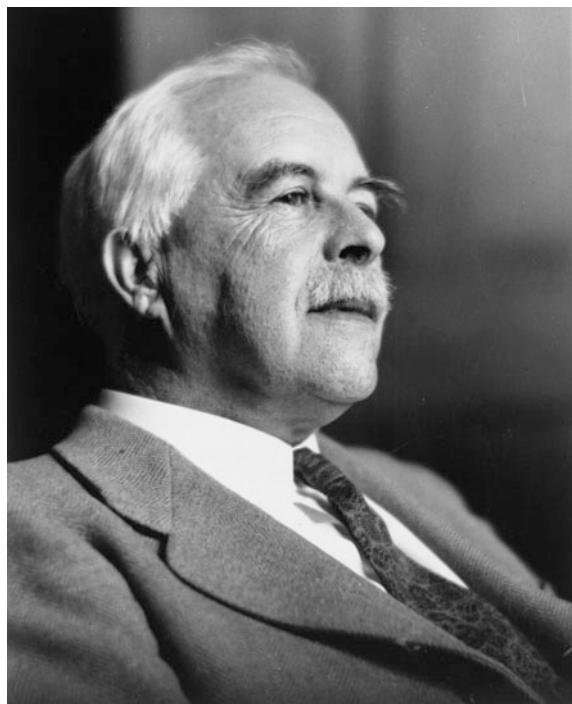
Boston and took a faculty position at MIT. In 1912 he joined the faculty of the University of California, Berkeley, where he spent the bulk of his scientific career.

Early on, Lewis was fascinated by thermodynamics, particularly as it pertained to chemical equilibrium, and wrote a number of papers on the then-nascent theory of relativity, proposing an alternate derivation of mass-energy equivalence from that of Einstein. In 1913, he was elected to the National Academy of Sciences (NAS).

Lewis is best known for his work on chemical bonding, particularly the notion of the covalent bond as outlined in his classic 1916 paper, “The Atom and the Molecule,” and a 1923 paper on the electron-pair theory of acid-base reactions. As early as 1902 he had been using rough sketches in his lecture notes, depicting cube-shaped atoms in which the corners represented possible positions of electrons. Eventually he realized that the electrons in an atom pair up around the nucleus, most commonly forming a tetrahedron, and that two bonding atoms could share paired electrons between them. These ideas later influenced the work of Linus Pauling, among others.

He was also well-known by coining the word “photon” to describe the unit of light in a December 18, 1926 letter to *Nature* (see preface of the present book for more details about the word “photon”-SGL). Technically, his understanding of the term was that it described a carrier of “radiant energy”—not a particle of light per se: “I therefore take the liberty of proposing for this hypothetical new atom, which is not light but plays an essential part in every process of radiation, the name photon.”

Fig. 15.1 G. N. Lewis.
Photograph by Johan
Hagemeyer, Bancroft Library,
University of California
Berkeley, courtesy AIP
Emilio Segrè Visual Archives



In that sense, his concept differed from Einstein's 1905 quantum theory of light, but "photon" came to be used to describe what Einstein originally termed Energiequanta.

Lewis has a street in his hometown of Weymouth named after him, as well as a wing of the local high school library. Berkeley's Lewis Hall is also named in his honor, and he received many professional accolades for his research. But the Nobel Prize in Chemistry eluded him. Some have speculated that his bitter rivalry with Nernst was partly to blame, with the latter using his position on the selection committee to block Lewis's nominations.

It was especially disappointing because one of Lewis's own students, Harold Urey, won the 1934 Nobel Prize in Chemistry for the discovery of deuterium—work which owed no small debt to Lewis's own contributions in using the Berkeley cyclotron to purify and characterize heavy water. Lewis resigned from the NAS that same year, possibly in a pique over this professional slight, although he may also have been frustrated by the failure of his own nominees to be elected to that August body. Another young physicist Lewis mentored, Glenn T. Seaborg, also went on to win the Nobel Prize in 1951. (Lewis did prefer to work with exceptional students, and their later success is, if nothing else, a testament to his excellent judgment).

For all his scientific accomplishments, Lewis met with a tragic end. In 1946, one of his graduate students came to the Berkeley lab and found Lewis's body under a workbench, apparently the victim of toxic fumes of liquid hydrogen cyanide. He had been working on an experiment using the substance, but a broken line caused the cyanide to leak into his laboratory workspace. Officially, his death was attributed to coronary artery disease, but rumors have persisted over the years that Lewis committed suicide.

The chemist had been struggling with depression, possibly stemming from his snubbing by the Nobel Prize committee despite a record in nominations over his lifetime (41 nominations—SGL).

On the day he died, Lewis had lunch with Irving Langmuir, who won the 1932 Nobel Prize in Chemistry for his work on surface chemistry and was in Berkeley to receive an honorary degree. Colleagues recalled Lewis came back from that lunch visibly moody. A few hours later, he was dead.

We likely will never know for certain whether Lewis took his own life, but his legacy lives on—not just in his scientific papers, and his two sons (both of whom became chemistry professors), but in the nomenclature he coined. Every time we use the word "photon," we pay some small homage to Gilbert Lewis.

- [1] Coffey, Patrick. *Cathedrals of Science: The Personalities and Rivalries That Made Modern Chemistry*. Oxford: Oxford University Press, 2008.
- [2] Lewis, Gilbert Newton: The conservation of photons, *Nature* 118 (2981), 874–875 (1926).
- [3] Lewis, Gilbert Newton. *Valence and the Structure of Atoms and Molecules*. New York: Chemical Catalog Co. Reprinted, New York: Dover, 1966.

15.2 Nonlinear Optical Effects in Phosphorescence and Absorption (1941)

This section overviews nonlinear optical effects reported in publication G. N. Lewis, D. Lipkin and T. T. Magel: Reversible photochemical processes in rigid media. A study of the phosphorescent state, *J. Am. Chem. Soc.* **63**, 3005–3018 (1941). It also contains a text excerpt and the figures reprinted with permission of the American Chemical Society from the original publication.

Studying the phosphorescence from a triplet state with a few seconds of a decay time of the fluorescein-boric acid phosphor, Lewis et al. have been able to observe saturation effects both in phosphorescence and absorption in this phosphor using the intense light from a high-pressure mercury arc (~ 1000 W).

Figure 15.2 shows that at low light intensities a phosphorescence signal is proportional to the intensity of excited light, but at higher intensity it tends toward a saturation. The abscissa shows the intensity of the exciting light and the ordinate represents intensity of phosphorescence, both in arbitrary units, a curve (1) is for the sample placed into an acetone bath at 18°C , and a curve (2)—into liquid air.

The General Electric Co., type A-H6, high-pressure mercury arc with a glass water-jacket and glass lens system was used. The intensity of the excitation light was varied by calibrated wire screens between the lamp and the phosphor.

For measuring the intensity of the phosphorescent light a Weston photronic cell (Model 594, Type 2) was used in conjunction with a lamp and a scale galvanometer, which linearity was tested before the measurements.

The phosphorescence (including its decay times) was investigated at various temperatures. For this purpose the phosphor was immersed into a dewar, containing a liquid at the desired temperature. Liquid hydrogen and liquid air were used for the two lowest temperatures, isopentane containing a small amount of ethanol was used up to -80°C , and from -80°C to room temperature—acetone.

The boric acid phosphors were prepared by heating crystalline boric acid mixed with the desired amount of fluorescein. When 90% of the combined water was driven off the boric acid, the resulting clear liquid was poured onto a metal plate and pressed with a second plate to form slabs.

Fig. 15.2 Saturation of phosphorescence of fluorescein in boric acid: (1) $+18^\circ\text{C}$; (2) -185°C

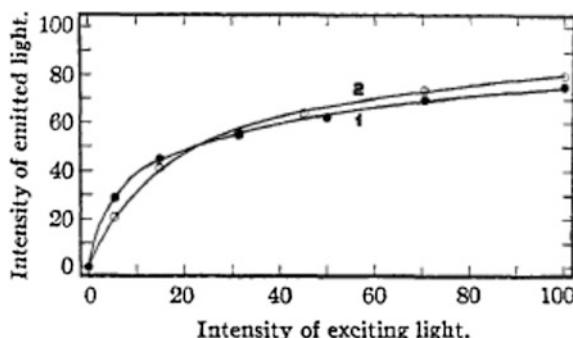


Fig. 15.3 Absorption spectrum of fluorescein in boric acid: (1) without illumination other than measuring light; (2) illuminated by high-pressure mercury arc (ϵ is molar extinction coefficient)

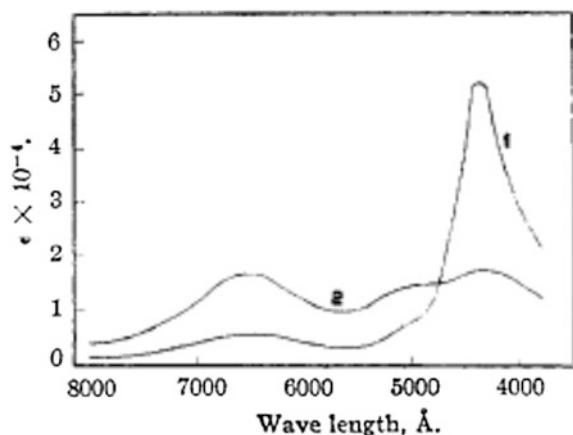
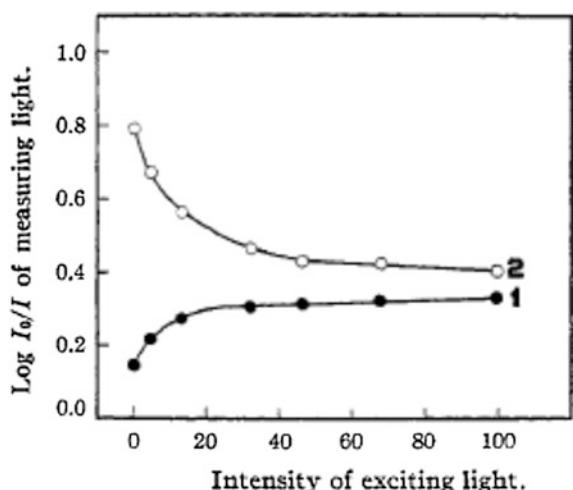


Figure 15.3 shows changing the absorption spectrum of fluorescein in boric acid from initial low intensity light illumination using a 500 W (General Electric Co. T-20) tungsten-filament projection lamp (curve 1) to high-intensity light illumination from the high-pressure mercury arc (curve 2). The lenses were used to focus the light from each source to the phosphor. The angle between low and high intensity light beams from two different light sources was about 30° at the sample. The sample thickness was ~ 1 mm.

The spectrometer for obtaining absorption spectra in these experiments contained a two-inch concave grating ruled with 14,437 lines per inch. A cesium oxide photoelectric cell (G-M Visitron, Type 75-AV) was used as a detector. This cell with the slit in front of it was moved across the spectrum by means of a screw. The entire assembly was housed in a thermostated box. The current from the photoelectric cell was amplified by means of an FP-54 vacuum tube amplifier and measured by means of a lamp and scale galvanometer.

Fig. 15.4 Effect of intensity of exciting light on the absorption of fluorescein in boric acid: (1) 6500 Å; (2) 4360 Å



The results of experiments on saturation of absorption with increasing light intensity are presented in Fig. 15.4. In this figure, $\log I_o/I$ of low intensity measuring light versus the intensity of the excitation light from the mercury arc is plotted, where I_o is the intensity of the incident measuring beam and I that of the transmitted beam. These measurements were carried out at wavelengths of 6500 and 4360 Å.

Curve 1 in Fig. 15.4 for 6500 Å shows the building up the substance which absorbs at this wavelength, namely, the phosphorescent state (see Fig. 15.3). As noticed by the authors, this experiment is comparable with data illustrated in Fig. 15.2, except that the zero of this curve is not for zero total illumination.

Setting the measuring apparatus at 4360 Å wavelength where only the “normal” molecules (which are not in the phosphorescence state), absorb (see Fig. 15.3), the authors obtained curve 2. This curve shows the gradual disappearance of the “normal” molecules, with increasing illumination by high-intensity excitation light.

The phosphorescence decay times measured in this paper at different temperatures reached the values of 2.5 s in some measurements.

Chapter 16

Nonlinear Optical Device (Icaroscope) at the Institute of Optics, University of Rochester During the Second World War



Carlos R. Stroud

Abstract This chapter contains a brief history of development of a nonlinear optical device (Icaroscope) in the University of Rochester during the Second World War based on saturation of luminescence (Sect. 16.1). The text of original publication by B. O'Brien is also included (Sect. 16.2).

16.1 A Brief History of Development of Icaroscope

Some excerpts of the text are reprinted from the book “*A Jewel in the Crown, 75th Anniversary Essays*”, The Institute of Optics, University of Rochester, C. R. Stroud, Jr., Ed., University of Rochester Press, 2004.

Work in nonlinear optics at The Institute of Optics can be traced back to the work of Brian O'Brien during the Second World War [1–4]. O'Brien had directed a number of research efforts aimed at helping the U.S. war effort. One such effort involved the development of phosphors that would emit visible light upon excitation by infrared radiation. This project was carried out under the immediate supervision of Franz Urbach (Fig. 16.1), former Austrian physicist immigrated to USA in 1939 after the “Anschluss”. Urbach continued his experimental work, begun in Vienna, on luminophores sensitive to infrared radiation. Later, after the war, Urbach worked in Phosphor Research Laboratories and Solid State Physics Laboratories of Eastman Kodak.

These phosphors led to the development of various infrared and low-light image converters that were crucial to the U.S. war effort. But an additional application used the nonlinear transfer characteristics of these phosphors [1–4]. A problem

C. R. Stroud (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road,
Wilmot Building, Rochester, NY 14627, USA
e-mail: stroud@optics.rochester.edu

facing the U.S. Navy was that the Japanese were making their bombing attacks on U.S. ships by approaching from the direction of the sun, thus effectively blinding those defending their ships. An imaging device was needed that would dramatically reduce the brightness of the sun while preserving the brightness of immediately adjacent objects. O'Brien and his associates solved this problem through use of saturation of the luminescence of their phosphors. Strong saturation of the luminescence efficiency occurred at those regions illuminated by the sun; little or no saturation occurred at the adjacent regions, allowing objects in these regions, such as enemy aircraft, to be clearly seen. The telescope formed an image on the phosphor screen which was then flipped so that the observer could view the screen in which the image of the sun itself was greatly attenuated compared with that of the attacking airplane. The screen was then flipped back to be exposed again. This process was repeated several times per second to follow the moving airplane.

This device became known as the Icaroscope, named after the Greek tragic hero Icarus (Fig. 16.2). Ironically, this same group also developed an instrument called the "Seebackascope" which allowed an attacking pilot to accurately align his airplane along a path coming exactly from the direction of the sun.

Fig. 16.1 Franz Urbach photographed by the light of the infrared phosphors that he developed. The photograph by George Burns appeared in the Saturday Evening Post magazine. [Reproduced with permission of University of Rochester, River Campus Libraries (University Archives)]

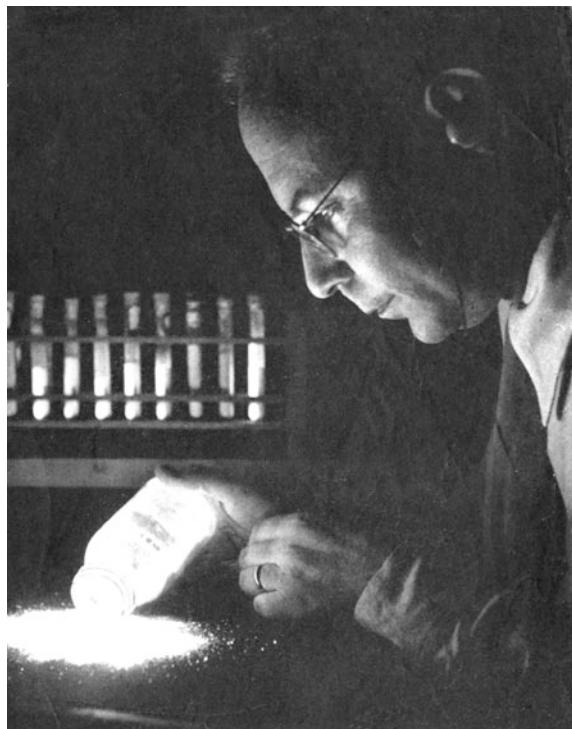




Fig. 16.2 Brian O'Brian using an Icaroscope is looking directly into the sun and suffering no discomfort. The Icaroscope provided effective defense from daylight dive-bombing attacks in the Second World War. [Reproduced with permission of University of Rochester, River Campus Libraries (University Archives)]

References

1. R.W. Boyd: History of research in nonlinear optics at the Institute of Optics, 293–300, in A Jewel in the Crown, 75th Anniversary Essays, The Institute of Optics, University of Rochester, C. R. Stroud, Jr., Ed., University of Rochester Press, 2004.
2. B. O'Brian: A story of science in night warfare, 65–75, *Ibid.*
3. C. Stroud: National defense research committee, 48–54, *Ibid.*
4. B. O'Brian: A brightness limiting phosphor telescope with selective action on the brighter portions of an image, *Journal of the Optical Society of America*, **36**, issue 12, 709 (1946).

16.2 Reprint of Original Publication About Icaroscope

From B. O'Brien, *Journal of the Optical Society of America* **36**, is. 12, 709 (1946)

Reprinted with permission from the Optical Society of America.

A brightness limiting phosphor telescope with selective action on the brighter portions of an image

A telescope-like device with selective brightness limiting properties has been developed, making use of the phenomenon of afterglow saturation in a phosphor screen. This has been named the Icaroscope, and makes possible simultaneous

observation of the sun and surrounding sky without being dazzled by the high brightness of the solar disk itself. An objective, especially designed to reduce troublesome flare, forms an image on a transparent phosphor screen exhibiting short period afterglow. The screen is viewed through an eyepiece and erecting system. Between objective and screen and between screen and eyepiece are two sector disk shutters rigidly mounted on the same motor shaft, but with the sector openings out of phase. Thus the screen is not visible while being illuminated, but is viewed a half-cycle later, after the illumination has ceased. This is repeated at the rate of 90 cycles per second so that viewing appears to be continuous, and the instrument is used as an ordinary telescope. The operation resembles the classical phosphoroscope with the important difference that the entire image of a scene is formed upon a stationary screen, and the phosphor is especially chosen for short afterglow with saturation. Saturation is very pronounced in the afterglow of certain sulphide phosphors. If an image of sun and surrounding sky is formed on such a phosphor, and the sample is then examined in the dark, the afterglow of the area which received the solar image will be only slightly greater than the surrounding region which was illuminated by the sky. The same effect is secured a few milliseconds after illumination by proper choice of phosphor, a silver activated zinc-cadmium sulphide having proven most satisfactory. As typical of icaroscope performance the rendition of all object brightnesses up to that of the sky at the sun's limb (object brightness approx. 50 lamberts) will give substantially normal contrast, yet the phosphor image of the solar disk (object brightness approx. 5×10^5 lamberts) will appear only about 20 times brighter than that of the surrounding sky. Special properties required in phosphor screen and transparent mount will be described.

Chapter 17

Measuring Sensitivity Threshold of the Human Eye to Feeble Light (Selig Hecht)



Svetlana G. Lukishova

Abstract This chapter is devoted to Selig Hecht (1892–1947) who was known for his substantial contributions to the study of photoreception and vision, especially for measuring sensitivity threshold of the human eye to feeble light. Section 17.1 contains Hecht's biography, in Sect. 17.2 the list of his papers related to vision is provided, and Sect. 17.3 contains a reprint of famous paper S. Hecht, S. Shlaer and M.H. Pirenne: Energy, Quanta and Vision, *J. Gen. Physiol.* 25, 819 (1942).

This chapter is dedicated to classical experiments by Selig Hecht et al. on measuring the sensitivity threshold of the human eye to feeble light. Born in Austria, Hecht was an outstanding American physiologist, who studied photochemistry in photoreceptor cells and the mechanisms of the visual threshold. For his work related to vision he was elected to the U.S. National Academy of Sciences (1944) and received the highest recognition of the Optical Society of America, the Frederick Ives Medal (1941).

Section 17.1 of this chapter is biographical and contains excerpts from a paper of B. O'Brian, H. Grundfest, and E. Smith: Selig Hecht (1892–1947), *Science*, **107**, Is. 2770, 105 (1948) reprinted with permission of the American Association for the Advancement of Science. See also *Selig Hecht: A Biographical Memoir* by George Wald, 81–100, National Academy of Sciences, Washington, D.C. (1991), <http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/hecht-selig.pdf>.

Section 17.2 presents the list of Selig Hecht's papers related to vision. In Sect. 17.3 his famous and widely cited paper is reproduced with permission by the Rockefeller University Press [S. Hecht, S. Shlaer and M. H. Pirenne: Energy, Quanta and Vision, *J. Gen. Physiol.* **25**, 819–840 (1942)]. Similar experiments on the sensitivity threshold of the human eye were carried out by Russian group of S.I. Vavilov (Wawilow). Chapter 14 of this book contains translation into English of the Brumberg and Vavilov paper of 1933 (Sect. 14.3) and a reproduction of the

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road,
Wilmot Building, Rochester, NY 14627, USA

e-mail: sluk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

1943 paper of Vavilov group (Sect. 14.5) on some of their measurements. In addition, Sect. 14.4 presents a translation of Vavilov's comments on the interpretation of his group's results by S. Hecht, S. Shlaer and M. H. Pirenne in their paper mentioned above and reproduced in Sect. 17.3.

Both the Hecht and Vavilov groups used classical, feeble light sources, and after adaptation to darkness, both groups estimated the human eye sensitivity threshold by fitting a Poisson distribution model of photons absorbed by the retina. The fitting curves' estimations in both groups' experiments showed that the human eye can see few photons. It is important that the eye sensitivity threshold depends on the slope of the curves in their fitting models which does not depend on absorption and scattering in the eye prior light absorption by the rods, most sensitive cells of retina. The human eye has $\sim 5\text{--}10\%$ of the total efficiency of light transmission from the cornea to its absorption in the periphery of retina where most rods are located (see page 10 of [1] for details). These experiments showed [although with some concerns regarding their accuracy from the modern point of view (see pages 14–16 of [1] for details)] that the individual rods can detect single photons, e.g., in Hecht's experiments with detection of 5–7 photons by the observer, the flashes of light were incident on the retina area containing about 500 rods, so in dark-adapted conditions individual rods must have detected the absorption of a single photon. In 1944 an independent experiment, similar to Vavilov and Hecht, was carried out by H. Van der Velden with a threshold of 1–2 photons [2]. In 1971 an important step was taken by Sakitt [3], who recognized some of the problems with the “did you see it or not?” experimental design used by Vavilov, Hecht, Van der Velden and the other early vision researchers. In [3] participants were asked to rate the brightness of feeble light flashes on a scale of 0–6, and at least one participant indicated the threshold of 1 photon. Later experiments with isolated rod photoreceptors (from the amphibian retina) with the same classical sources provided direct evidence that the rods produce electrical signals in response to faint flashes of light, with the smallest signals corresponding to single photons [4].

Recent applications of heralded single-photon sources based on spontaneous parametric downconversion (SPDC) to the visual perception by the human eye [1, 5–9] permitted to avoid some restrictions of earlier measurements with conventional light sources with Poisson statistics [1, 9]. As accepted by the scientific community [9], the best proof to date that the whole visual system of the human eye can detect a single photon was provided in the experiment [7] with a heralded SPDC single-photon source. In [8], using a SPDC source, temporal integration in the visual system at the few-photon level was studied. Much longer integration times (650 ms in average) were obtained in [8] than have been suggested earlier (100 ms). With a SPDC source the quantum efficiency of an isolated rod photoreceptor cell was measured to be about $(29 \pm 4.7)\%$ for *Xenopus laevis* toads [10]. A single rod photoreceptor (*Xenopus laevis*) was used to measure photon statistics of different light sources (coherent and pseudothermal) [11]. Chapter 6 of the current book discusses the role of photon statistics in visual perception and overviews papers in this direction.

The success of new experiments on visual perception with SPDC sources enables the studies of quantum effects by the human eye as a detector. For instance, in [9] (see also other references in it, e.g. [12]) two experiments are proposed: (1) to determine whether humans perceive any difference between a photon in a superposition state and a classical mixed state, and (2) to use a human observer as a detector in a Bell test. See also [13, 14] on proposals of using the human eye in a Bell test¹.

- [1] R. Holmes: Testing the limits of human vision with quantum states of light, *Ph.D dissertation*, University of Illinois at Urbana-Champaign, 2017.
- [2] H. Van der Velden: The number of quanta necessary for the perception of light of the human eye, *Ophthalmologica*, **111** (6), 321–331 (1946), originally published in 1944 in Dutch.
- [3] B. Sakitt: Counting every quantum, *The Journal of Physiology*, **223**(1), 131–150 (1972)
- [4] F. Rieke, D.A. Baylor: Single photon detection by rod cells of the retina. *Reviews of Modern Physics*. **70**, 1027–1036 (1998).
- [5] R. Holmes, B.G. Christensen, W. Street, R.F. Wang, and P.G. Kwiat: Determining the lower limit of human vision using a single photon source, in *Conference on Lasers and Electro-Optics*, paper QTU1E.8, Optical Society of America (2012).
- [6] R. Holmes, B.G. Christensen, R.F. Wang, and P.G. Kwiat: Testing the limits of human vision with single photons, in *Frontiers in Optics*, paper FTU5B.5, Optical Society of America (2015).
- [7] J.N. Tinsley, M.I. Molodtsov, R. Prevedel, D. Wartmann, J. Espigulé-Pons, M. Lauwers, and A. Vaziri: Direct detection of a single photon by humans, *Nature Communications* **7**, 12172 (July 2016).
- [8] R. Holmes, M. Victora, R.F. Wang, and P.G. Kwiat: Measuring temporal summation in visual detection with a single-photon source, *Vision Research* **140**, 33–43 (2017).
- [9] R.M. Holmes, M.M. Victora, R.F. Wang, and P.G. Kwiat: Testing the limits of human vision with quantum states of light: past, present, and future experiments, *Proc. SPIE* **10659**, *Advanced Photon Counting Techniques XII*, paper 1065903 (14 May 2018).
- [10] N.M. Phan, M.F. Cheng, D.A. Bessarab, and L.A. Krivitsky: Interaction of fixed number of photons with retinal rod cells, *Phys. Rev. Lett.* **112**, 213601 (May 2014).
- [11] N. Sim, M.F. Cheng, D.A. Bessarab, C.M. Jones, and L.A. Krivitsky: Measurement of photon statistics with live photoreceptor cells, *Phys. Rev. Lett.* **109**, 113601 (Sep 2012).
- [12] G. Ghirardi: Quantum superpositions and definite perceptions: envisaging new feasible experimental tests, *Physics Letters A* **262**, 1–14 (Oct 1999).
- [13] N. Brunner, C. Branciard, and N. Gisin: Possible entanglement detection with the naked eye. *Phys. Rev. A* **78**, 052110 (2008).

¹The author thanks R. Holmes for her comments.

- [14] V. Caprara Vivoli, P. Sekatski, and N. Sangouard: What does it take to detect entanglement with the human eye? *Optica* **3**, 473–476 (May 2016).

17.1 Selig Hecht: Biography

Excerpts from Brian O'Brian, Harry Grundfest, Emil Smith: Selig Hecht (1892–1947), *Science*, **107**, Is. 2770, 105 (1948) are reprinted with permission of the American Association for the Advancement of Science.

Selig Hecht (Fig. 17.1) was born in Glogow, Austria, on February 8, 1892. He received his education in the public school system of New York City, obtaining his B.Sc. degree from the College of the City of New York in 1913. In 1913–14 he was a pharmacologist with the U.S. Department of Agriculture, going thence to Harvard for graduate training under G. H. Parker.

From 1917, when he obtained his Ph.D., to 1921, Hecht was assistant professor of physiology at the College of Medicine, Creighton University. For the next 5 years, as a National Research Council Fellow in chemistry, he worked at Liverpool and Cambridge Universities in England, at the Naples Zoological Laboratory, at the University of Berlin, and at the Harvard Medical School. During these years he acquired a wide international acquaintance and a circle of friends which he maintained and cherished.

Going to Columbia University in 1926, Hecht organized the Laboratory of Biophysics and in 1928 became professor of biophysics. He set and maintained an extraordinarily high standard for the work of his laboratory, which he developed as one of the productive research centers of the country. Nearly a score of students whom he trained branched out in widely divergent aspects of general physiology and are actively engaged in research in biology, biochemistry, physiology, biophysics, and ophthalmology.

Except for his earliest scientific explorations, Selig Hecht worked entirely on the problems of photoreception and vision, and few investigators have ever become such masters of their field. He pioneered in applying physicochemical principles to sensory physiology, using these principles lucidly and logically to demonstrate the existence of distinct chemical reactions in photoreception, to formulate quantitatively his concept of the sensory state, and to extend its applications to many visual phenomena.

Hecht's best-known work was the determination of minimal quantal requirements at the threshold of vision for the human eye. Characteristically, the approach was made by two independent methods which gave the same result.

Hecht's wide knowledge of human vision was devoted during the war to problems of military importance. He served on the Subcommittee of Visual Problems, Division of Medical Sciences, National Research Council, and on the Army-Navy Office of Scientific Research and Development Vision Committee, and was official investigator for several contracts with the Army, Navy, and Air Forces.

Fig. 17.1 Selig Hecht. AIP Emilio Segrè Visual Archives, Physics Today Collection



For his work on vision Hecht received the Frederick Ives Medal of the Optical Society of America in 1941. He was elected to the U.S. National Academy of Sciences in 1944. He was a member of many scientific societies which reflected his interests in physics (Optical Society), in physiology (American Physiological Society), and in biology (American Society of Naturalists).

Selig Hecht did not confine his interests to science alone. He was a musician and painter, a devotee of literature and the arts. These enthusiasms, as well as his interest in the political and social currents of the times, he shared with his students, imparting to them his sense of the connection of science, the arts, and society.

An important facet of Hecht's personality was his awareness of the social aspects of science. He was authoritatively acquainted with current views on, and experiments in, improving science teaching in the colleges and was a participant in his own university's studies of this subject. He devoted considerable time and effort transmitting to very wide circles his love for science. His lecture courses at the New School for Social Research were of such caliber that Dr. Alvin Johnson, president emeritus of the New School, characterized Selig Hecht as "absolutely the best teacher I have ever known." Hecht's last book (*Explaining the atom*. New York: Viking, 1947), on atomic energy for the general public, has received wide praise for its clear and very readable approach. This book was a result of his interest and

experience in popular science education, and of his concern for a world facing the task of controlling atomic energy by converting it from a weapon of great destruction to a powerful tool of civilization. Because of this concern, Hecht was an active member of the Emergency Committee of Atomic Scientists—the only one of this small group who was not a nuclear scientist.

All who came in contact with Selig Hecht immediately grasped one of his most important attributes, his joyousness. Hecht enjoyed everything he did, in science, in his painting, in living with his family, his friends, and his colleagues. This joy filled him with enthusiasm for his every activity. It was an essential ingredient in his excellence as a scientist, lecturer, teacher, and writer. To the task in hand he gave everything of himself. He loved intellectual exercise, and, when the occasion seemed to demand it, intellectual combat. Yet this joy made Hecht friendly, sympathetic, and always approachable.

17.2 List of Selig Hecht Papers Related to Vision

Most of these references is excerpt from the Journal of the Optical Society of America **32**, 40–41 (1942).

1. "The photic sensitivity of *Ciona Intestinalis*," J. Gen. Physiol. 1, 147 (1918).
2. "Sensory equilibrium and dark adaptation in *Mya Arenaria*," J. Gen. Physiol. 1, 545 (1919).
3. "The nature of the latent period in the photic response of *Mya Arenaria*," J. Gen. Physiol. 1, 657 (1919).
4. "The effect of temperature on the latent period in the photic response of *Mya Arenaria*," J. Gen. Physiol. 1, 667 (1919).
5. "The photochemical nature of the photosensory process," J. Gen. Physiol. 2, 229 (1920).
6. "Intensity and the process of photoreception," J. Gen. Physiol. 2, 337 (1920).
7. "Human retinal adaptation," Proc. Nat. Acad. Sci. 6, 112 (1920).
8. "The dark adaptation of the human eye," J. Gen. Physiol. 2, 499 (1920).
9. "Photochemistry of visual purple. I. The kinetics of the decomposition of visual purple by light," J. Gen. Physiol. 3, 1 (1920).
10. "Photochemistry of visual purple. II. The effect of temperature on the bleaching of visual purple by light," J. Gen. Physiol. 3, 285 (1921).
11. "Time and intensity in photosensory stimulation," J. Gen. Physiol. 3, 367 (1921).
12. "The relation between the wave-length of light and its effect on the photosensory process," J. Gen. Physiol. 3, 375 (1921).
13. "The photochemistry of the sensitivity of animals to light," Science 53, 347 (1921).
14. "The photochemistry of the sensitivity of animals to light," J. Opt. Soc. Am. 5, 227 (1921).

15. "The nature of foveal dark adaptation," J. Gen. Physiol. 4, 113 (1921).
16. "The visibility of monochromatic radiation and the absorption spectrum of visual purple" (with R. E. Williams), J. Gen. Physiol. 5, 1 (1922).
17. "Sensory adaptation and the stationary state," J. Gen. Physiol. 5, 555 (1923).
18. "Intensity discrimination and the stationary state," J. Gen. Physiol. 6, 355 (1924).
19. "Photochemistry of visual purple. III. The relation between the intensity of light and the rate of bleaching of visual purple," J. Gen. Physiol. 6, 731 (1924).
20. "The visibility of the spectrum," J. Opt. Soc. Am. 9, 211 (1924).
21. "The visual discrimination of intensity and the Weber-Fechner law," J. Gen. Physiol. 7, 235 (1924).
22. "Zur Photochemie des Sehens," Naturwiss. 13, 66 (1925).
23. "Zur Theorie des Sehens," Naturwiss. 13, 660 (1925).
24. "The general physiology of vision," Am. J. Physiol. Opt. (Amer. Opt. Co.) 6, 303 (1925).
25. "The effect of exposure period and temperature on thephotosensory process in Ciona," J. Gen. Physiol. 8, 291 (1926).
26. "A quantitative basis for visual acuity and intensity discrimination," Skand. Arch. Physiol. 49, 146 (1926).
27. "The kinetics of dark adaptation," J. Gen. Physiol. 10, 781 (1927).
28. "A quantitative basis for the relation between visual acuity and illumination," Proc. Nat. Acad. Sci. 13, 569 (1927).
29. "The relation between visual acuity and illumination," J. Gen. Physiol. 11, 255 (1928).
30. "On the binocular fusion of colors and its relation to theories of color vision," Proc. Nat. Acad. Sci. 14, 237 (1928).
31. "The influence of temperature on the photosensory latent period," J. Gen. Physiol. 11, 649 (1928).
32. "The relation of time, intensity, and wavelength in the photosensory system of Pholas," J. Gen. Physiol. 11, 657 (1928).
33. "Visual acuity and illumination," Arch. of Ophthal. 57, 564 (1928).
34. "The visual acuity of the bee and its relation to illumination" (with E. Wolf), Proc. Nat. Acad. Sci. 15, 178 (1929).
35. "The nature of the sensitivity of animals to light," J. Opt. Soc. Am. 18, 264 (1929).
36. "The visual acuity of the honey bee" (with E. Wolf), J. Gen. Physiol. 12, 727 (1929).
37. *The Foundations of Experimental Psychology* (1929), Chapter V, 216. Vision: II. The nature of the photoreceptor process.
38. "The visual acuity of insects" (with E. Wolf and G. Wald), Am. J. Physiol. 90, 381 (1929).
39. "Eine Grundlage fur die Beziehung zwischen Sehschärfe und Beleuchtung," Naturwiss. 18, 233 (1930).

40. "The intensity factor in vision and radiation," Am. Naturalist 64, 193 (1930).
41. "The development of Thomas Young's theory of color vision," J. Opt. Soc. Am. 20, 231 (1930).
42. "The retinal processes concerned with visual acuity and color vision," Bull. No. 4 of the Howe Laboratory of Ophthalmology, Harvard Medical School (Aug. 1931).
43. "Die physikalische Chemie und die Physiologie des Sehaktes," Ergeb. der Physiol. 32, 243 (1931).
44. "The interrelation of various aspects of color vision," J. Opt. Soc. Am. 21, 615 (1931).
45. "Intermittent stimulation by light. I. The validity of Talbot's law for Mya" (with E. Wolf), J. Gen. Physiol. 15, 369 (1932).
46. "The dark adaptation of different retinal areas" (with G. Wald and C. Haig), Am. J. Physiol. 101, 52 (1932).
47. "A quantitative formulation of colour-vision," Phys. Soc. London, joint discussion on "Vision," 126 (1932).
48. "The influence of intensity, color and retinal location on the fusion frequency of intermittent illumination" (with C. D. Verrijp), Proc. Nat. Acad. Sci. 19, 522 (1933).
49. "Intermittent stimulation by light. II. The measurement of critical fusion frequency for the human eye" (with S. Shlaer and C. D. Verrijp), J. Gen. Physiol. 17, 237 (1933).
50. "Intermittent stimulation by light. III. The relation between intensity and critical fusion-frequency for different retinal locations" (with C. D. Verrijp), J. Gen. Physiol. 17, 251 (1933).
51. "Intermittent stimulation by light. IV. A theoretical interpretation of the quantitative data of flicker" (with C. D. Verrijp), J. Gen. Physiol. 17, 269 (1933).
52. "The dark adaptation of various retinal areas" (with G. Wald and C. Haig), Arch. di Sci. Biol. 18, 170 (1933).
53. "The influence of intensity on the visual functions of drosophila" (with G. Wald), Proc. Nat. Acad. Sci. 19, 964 (1933).
54. "The visual acuity and intensity discrimination of drosophila" (with G. Wald), J. Gen. Physiol. 17, 517 (1934).
55. A Handbook of General Experimental Psychology (1934), Chapter 14, p. 704. Vision: II. The nature of the photoreceptor process.
56. "Anomalies in the absorption spectrum of visual purple" (with A. M. Chase), Proc. Nat. Acad. Sci. 20, 238 (1934).
57. "A theoretical basis for intensity discrimination in vision," Proc. Nat. Acad. Sci. 20, 644 (1934).
58. "A theory of visual intensity discrimination," J. Gen. Physiol. 18, 767 (1935).
59. "Intensity discrimination," Cold Spring Harbor Symposia on Quantitative Biology 3, 230 (1935).

60. "Intermittent light stimulation and the duality theory of vision" (with S. Shlaer and E. L. Smith), *ibid.* 3, 237 (1935).
61. "The basis for wavelength discrimination by the color-blind" (with S. Shlaer), *Sech. J. Physiol. USSR* 21, 367 (1935).
62. "The influence of color and area on the fusion frequency of intermittent stimulation" (with E. L. Smith and S. Shlaer), *Sech. J. Physiol. USSR* 21, 363 (1935).
63. "The dark adaptation of retinal fields of different size and location" (with C. Haig and G. Wald), *J. Gen. Physiol.* 19, 321 (1935).
64. "Intermittent stimulation by light. V. The relation between intensity and critical frequency for different parts of the spectrum" (with S. Shlaer), *J. Gen. Physiol.* 19, 965 (1936).
65. "Intermittent stimulation by light. VI. Area and the relation between critical frequency and intensity" (with E. L. Smith), *J. Gen. Physiol.* 19, 979 (1936).
66. "La Nature de la Discrimination d'Intensité dans la Vision et la Photoreception," *J. de Psychologie*, 33, 161 (1936).
67. "The color vision of dichromats" (with S. Shlaer). I. "Wavelength discrimination, brightness distribution, and color mixture." *J. Gen. Physiol.* 20, 57 (1936).
68. II. "Saturation as the basis for wavelength discrimination and color mixture." *J. Gen. Physiol.* 20, 83.
69. "Dark adaptation as influenced by intensity of light adaptation" (with C. Haig), *J. Opt. Soc. Am.* 26, 304 (1936).
70. "The regeneration of visual purple in solution" (with A. M. Chase, S. Shlaer, C. Haig), *Science* 84, 331 (1936).
71. "Intensity discrimination and its relation to the adaptation of the eye," *J. Physiol.* 86, 15 (1936).
72. "The influence of light adaptation on subsequent dark adaptation of the eye" (with C. Haig and A. M. Chase), *J. Gen. Physiol.* 20, 831 (1937).
73. "The instantaneous visual threshold after light adaptation," *Proc. Nat. Acad. Sci.* 23, 227 (1937).
74. "Visual intensity discrimination in different parts of the spectrum" (with J. C. Peskin and M. Patt), *Am. J. Physiol.* 119, 330 (1937).
75. "Rods, cones and the chemical basis of vision," *Physiol. Reviews* 17, 239 (1937).
76. "The diffusion coefficient and molecular size of visual purple" (with A. M. Chase and S. Shlaer), *Science* 85, 567 (1937).
77. "The nature of the visual process," *Bull. N. Y. Acad. Med.* 14, 21 (1938).
78. "The photochemical basis of vision," *J. App. Phys.* 9, 156 (1938).
79. "The visual functions of a completely color-blind person" (with S. Shlaer, E. L. Smith, C. Haig, and J. C. Peskin), *Am. J. Physiol.* 123, 94 (1938).
80. "The sedimentation constant of visual purple" (with E. G. Pickels), *Proc. Nat. Acad. Sci.* 24, 172 (1938).

81. "La base chimique et structural de la vision," Exposes de Biophysique, 583, 100 pp. (1938).
82. "Vitamin A and rod-cone dark adaptation in cirrhosis of the liver" (with C. Haig and A. J. Patek), *Science* 87, 534 (1938).
83. "An adaptometer for measuring human dark adaptation," *J. Opt. Soc. Am.* 28, 269 (1938).
84. "Rod-cone dark adaptation and Vitamin A" (with J. Mandelbaum), *Science* 88, 219 (1938).
85. "Intensity discrimination in the human eye. II. The relation between $\Delta I/I$ and intensity for different parts of the spectrum" (with J. C. Peskin and M. Patt), *J. Gen. Physiol.* 22, 7 (1938).
86. "The visibility of single lines at various illuminations and the retinal basis of visual resolution" (with Esther U. Mintz), *J. Gen. Physiol.* 22, 593 (1939).
87. "Diagnosis and incidence of Vitamin A deficiencies," *J. Conn. State Med. Soc.* 3, 223 (1939).
88. "The relation between Vitamin A and dark adaptation" (with J. Mandelbaum), *J. Am. Med. Assoc.* 112, 1910 (1939).
89. "The sensibility of the nocturnal long-eared owl in the spectrum" (with M. H. Pirenne), *J. Gen. Physiol.* 23, 709 (1940).
90. "Dark adaptation and experimental human Vitamin A deficiency" (with J. Mandelbaum), *Am. J. Physiol.* 130, 651 (1940).
91. The Chemistry of Vision, "Modern trends in ophthalmology," Chapter 31, p. 328 (1940).
92. "Energy at the threshold of vision" (with S. Shlaer and M. H. Pirenne), *Science* 93, 585 (1941).
93. S. Hecht, S. Shlaer and M.H. Pirenne: Energy, quanta and vision, *J. Gen. Physiol.*, 25, 819–840 (1942).
94. S. Hecht: The quantum relations of vision, *Journal of the Optical Society of America* 32, 42–49 (1942).
95. S. Hecht: Energy and vision (Sigma Xi National Lecture), *American Scientist* 32, N 3, 159–177 (1944).
96. S. Hecht, S. Shlaer, E.L. Smith, C. Haig, and J.C. Peskin: The visual functions of the complete colorblind, *J. Gen. Physiol.* 31, 459–472 (1948).

17.3 Energy, Quanta and Vision (Original Paper)

Reproduced with permission by the Rockefeller University Press from S. Hecht, S. Shlaer and M.H. Pirenne, *J. Gen. Physiol.* 25, 819–840 (1942).

ENERGY, QUANTA, AND VISION*

BY SELIG HECHT, SIMON SHLAER, AND MAURICE HENRI PIRENNE†

(*From the Laboratory of Biophysics, Columbia University, New York*)

(Received for publication, March 30, 1942)

I

Threshold Energies for Vision

The minimum energy required to produce a visual effect achieves its significance by virtue of the quantum nature of light. Like all radiation, light is emitted and absorbed in discrete units or quanta, whose energy content is equal to its frequency ν multiplied by Planck's constant h . At the threshold of vision these quanta are used for the photodecomposition of visual purple, and in conformity with Einstein's equivalence law each absorbed quantum transforms one molecule of visual purple (Dartnall, Goodeve, and Lythgoe, 1938). Since even the earliest measurements show that only a small number of quanta is required for a threshold stimulus, it follows that only a small number of primary molecular transformations is enough to supply the initial impetus for a visual act. The precise number of these molecular changes becomes of obvious importance in understanding the visual receptor process, and it is this which has led us to the present investigation.

The first measurements of the energy at the visual threshold were made by Langley (1889) with the bolometer he invented for such purposes (Langley, 1881). He found the energy to be 3×10^{-9} ergs for light of 550 m μ . Langley worked before the physiology of vision was understood, so that he used the wrong light and took none of the precautions now known to be necessary; even so, his results are too high only by a factor of 10.

In the fifty years since Langley there have been eleven efforts to redetermine the minimum energy for vision. We have carefully studied all these accounts and have done our best to evaluate the measurements. Unfortunately, many of them contain serious errors which invalidate them. Most of them involved no direct energy determinations; instead, the investigators relied on previously measured energy distributions in standard sources and made elaborate computations from them. Only a few can be considered as reliable.

After Langley, the earliest paper is by Grijns and Noyons (1905). Their data differ widely from all other measurements and cannot be accepted even

* A preliminary report of these measurements was published in *Science* (Hecht, Shlaer, and Pirenne, 1941), and presented to the Optical Society in October, 1941 (Hecht, 1942).

† Fellow of the Belgian American Educational Foundation.

though it is hard to discover their precise errors because the description is too obscure. Zwaardemaker (1905), in whose laboratory their measurements were made, reports some of his own rough determinations, which turn out to be near Langley's. Neither Grijns and Noyons nor Zwaardemaker actually measured the energies involved, but relied on Ångström's (1903) determinations of the energy distribution in the Hefner lamp.

The best of the early efforts is by von Kries and Eyster (1907); and though the results involve many calculations, they come very close to the most careful of modern measurements. Von Kries and Eyster made no direct energy determinations; they measured brightnesses, durations, and areas. The conversion of these factors into final energies requires skill and care in the evaluation of absorptions, reflections, lens factors, and the like, and it is gratifying to see the admirable way in which von Kries accomplished this task.

TABLE I
Minimum Energy for Vision

Wavelength <i>mμ</i>	Energy <i>ergs</i>	No. of quanta	Source
505	$0.66\text{--}1.17 \times 10^{-10}$	17-30*	Chariton and Lea (1929)
507	$1.3\text{--}2.6 \times 10^{-10}$	34-68	von Kries and Eyster (1907)
530	$1.5\text{--}3.3 \times 10^{-10}$	40-90	Barnes and Czerny (1932)

* For inexperienced observers.

Computations from star magnitudes were made by Ives (1916) and by Russell (1917). However, neither they nor Reeves (1917) and Buisson (1917), who both reproduced star observations in the laboratory, employed the best physiological conditions for the measurements. Moreover, none of them took consideration of the different luminosity curves for rod vision and cone vision, and used the latter as standard in the computations.

Direct energy measurements were made by du Noüy (1921), but his work involves serious physical errors, and his results are too low by a factor of more than 100—so low indeed as to seem impossible.

The most recent determinations are by Chariton and Lea (1929), by Wentworth (1930), and by Barnes and Czerny (1932), all of whom agree in the order of magnitude of their results. Wentworth's exposures were too long to yield minimal values; otherwise her work is excellent. She measured the energies involved, which Barnes and Czerny also did, but not as directly.

From these twelve researches, we have chosen the three sets of measurements which are free from what can now be recognized as obvious error. These are given in Table I. Even though they differ by a factor of about 3, these data can be considered as roughly confirming one another. However,

S. HECHT, S. SHLAER, AND M. H. PIRENNE

821

since for our purposes a factor of 3 cannot be ignored, we undertook to make the measurements again, but under the best physical and physiological conditions.

II

Visual Conditions

The circumstances which will yield the maximum retinal sensibility have been adequately known for years. They involve dark adaptation, peripheral vision, small test fields, short exposures, and selected portions of the spectrum.

Complete dark adaptation means a stay of at least 30 minutes in the dark before measurements can be begun (Piper, 1903; Hecht, Haig, and Chase, 1937). After thorough dark adaptation the periphery of the retina is much more sensitive than its center. The greatest density of rod elements begins at about 18° out ($\text{\O}sterberg$, 1935), and exploration shows that between 20 and 30° from the center there is a region of maximum sensibility to light (Wentworth, 1930). The variation within this region is not large, and for convenience we chose a retinal area situated 20° temporally on the horizontal axis.

In visual threshold measurements it has been established that the larger the test area, the smaller need the intensity be for its recognition (*cf.* summary by Wald, 1938 *a*). This reciprocal relation is exact only for small areas. Our preliminary experiments, as well as the work of other investigators, show a minimum for the product of area and intensity for fields of the order of 10 minutes diameter. We therefore chose a circular retinal area of 10 minutes diameter for the test field.

The energy required to pass over the visual threshold involves an approximately reciprocal relationship between intensity and time of exposure. For exposures shorter than 0.01 second, the reciprocal relation holds perfectly (Graham and Margaria, 1935). To be sure of falling within this most efficient range, our exposures were 0.001 second long.

Finally, from the measurements of the scotopic luminosity curve (Hecht and Williams, 1922), it is known that for dim vision the eye is most sensitive to a wavelength of $510 \text{ m}\mu$, and this is the light which we used for making the measurements.

III

Apparatus and Calibrations

The physical arrangements may be seen in Fig. 1. The light source L is a ribbon filament lamp run on constant current obtained from storage cells and measured potentiometrically. By means of a lens, it is focussed on the slit of a double monochromator M_1M_2 and finally on the artificial pupil P . The subject, who sits in a dark cabinet in the dark room, has his head in a fixed position by keeping his teeth in a "bite" or hard impression of his upper jaw. He has his left eye next to the pupil P , and on looking at the red fixation point FP he sees the field lens FL . The light intensity of this

uniformly illuminated field is varied in large steps by the neutral filters F , and in a gradual way by the neutral wedge and balancer W . The size of the field is controlled by the diaphragms D . Its exposure is fixed by the shutter S , and is initiated by the subject.

For the record it is necessary to describe the apparatus and calibrations in detail. The double monochromator is made of two individual constant deviation monochromators, M_1 and M_2 , which are arranged for zero dispersion by means of the reversing prism RP . In this way, all the light passes through an equal thickness of glass, and assures a uniform brightness of the field lens FL . The exit slit of M_1 has been removed, and the entrance slit of M_2 serves as the middle slit of the combined double monochromator. The entrance and exit slits of the combination are kept at 1.2 mm., which corresponds to a band width of $10 \text{ m}\mu$ centered at $510 \text{ m}\mu$. The middle slit, before which the shutter is placed, is kept at 0.1 mm.

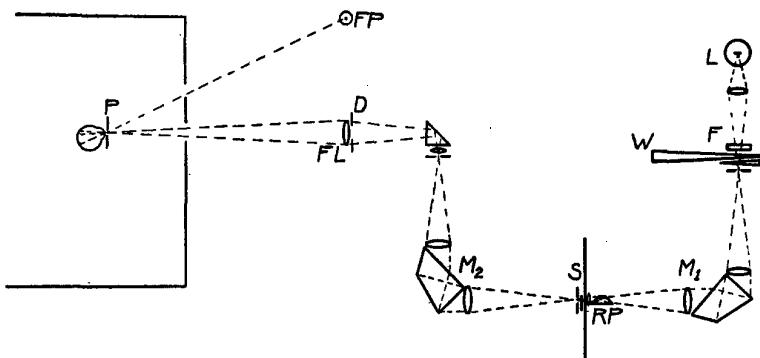


FIG. 1. Optical system for measuring minimum energies necessary for vision. The eye at the pupil P fixates the red point FP and observes the test field formed by the lens FL and the diaphragm D . The light for this field comes from the lamp L through the neutral filter F and wedge W , through the double monochromator M_1M_2 and is controlled by the shutter S .

The field lens FL magnifies the exit slit by a factor of 2, and thus yields an image of it 2.4 mm. wide and over 10 mm. high at the pupil P . The image is sufficient to cover uniformly not only the pupil P , but also the linear thermopile used for the energy calibration. The pupil mount at P and the field lens FL are connected by a carefully diaphragmed and blackened tube. The 2 mm. circular pupil P used for the visual measurements can be replaced by a slit 2 mm. wide and 10 mm. high behind which is the receiving surface of the thermopile for energy measurements.

S is a precision shutter made of two parts. One part is a thin circular aluminum disc with a small sector of 10.8° removed and properly balanced. It is run at 1800 R.P.M. by means of a synchronous motor, and therefore permits light to pass through the middle slit for $1/1000$ second during each revolution. The other part is a polar relay shutter, which, by means of a phasing commutator on the shaft of the synchronous motor, is opened for only one passage of the rotating disc aperture whenever the subject releases a push button.

S. HECHT, S. SHLAER, AND M. H. PIRENNE

823

The essentials of the shutter are shown in diagrammatic detail in Fig. 2. On the same shaft with the disc there is mounted a commutator having a "live" sector, which together with the brush occupies somewhat less than 90°. Two brushes are arranged on this commutator 90° apart, and are so phased with the A.C. line voltage that one of these brushes receives only a positive impulse while the other receives only a negative impulse. These impulses control a polar relay PR_2 , which then actuates a pair of single pole, double throw micro switches, MS_1 and MS_2 . These are arranged with their springs in opposition in such a manner that the switches are in equilibrium

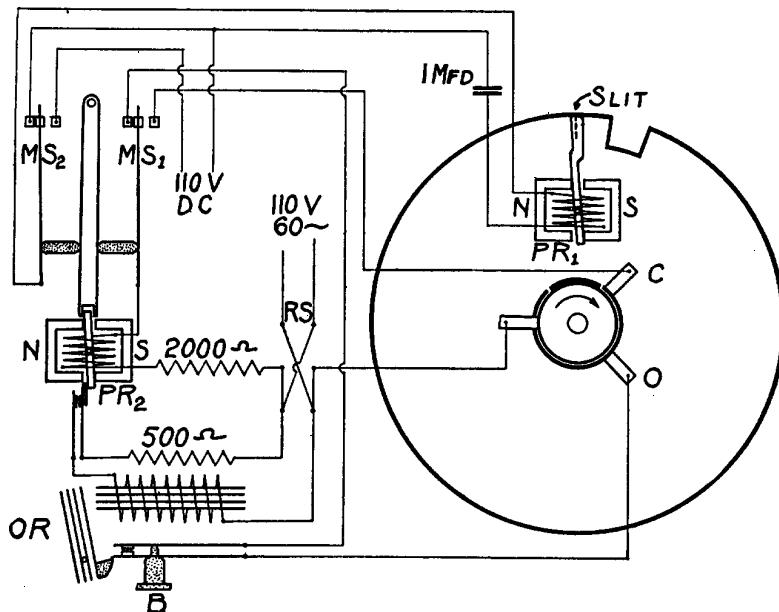


FIG. 2. Shutter for obtaining a single exposure of 1/1000 second. The details are described in the text.

at either of their two positions, and require but a small force and movement to kick them over to their other positions. Micro switch MS_1 is in series with the winding of PR_2 , and in one position connects with the opening brush O and in the other position with closing brush C . The other micro switch, MS_2 , charges and discharges a 1 μ fd. condenser from the 110 volt, d.c. line through the polar relay PR_1 . These impulses in and out of the condenser actuate PR_1 whose armature movement then uncovers and covers the middle slit.

The operation is seen by following a single cycle of operation of circuit and shutter. Fig. 2 shows the apparatus during its rest or closed period. The 110 volt, 60 cycle power enters through a pole-reversing switch, RS , to the neutral brush on the commutator. The impulse through the closing brush C is blocked, since it enters an open

contact in MS_1 . The impulse going to the opening brush O is blocked at the secondary contacts of the overload relay OR , the push button of which, B , is controlled by the observer. When B is released so that contact is made, the next impulse which leaves the opening brush O goes through the left hand contact of MS_1 , through the winding of PR_2 , and through a 2,000 ohm resistance to the other side of the power line. This impulse through PR_2 is adequate to throw its armature to the other position, thus switching over both MS_1 and MS_2 , and closing the power circuit through the primary of the overload relay OR . The activation of OR closes its armature, whose movement opens the secondary contacts attached to it, thereby breaking the circuit from the opening brush O so that the cycle does not repeat itself. The switching of MS_2 to its right contact charges the 1 μ fd. condenser through PR_1 , which moves its armature and thereby exposes the slit. The switching of MS_1 to its right contact sets the circuit for the very next impulse through the closing brush C to PR_2 . This closing impulse comes exactly three half-cycles or 3/120 second after the original opening impulse, and causes PR_2 to return to its original position. Now MS_2 discharges the 1 μ fd. condenser, which actuates PR_1 so that its armature moves to cover the slit and terminate the cycle.

The pole-reversing switch RS enables one to select the correct polarity for the operation of this circuit. It needs to be set only at the beginning of an experiment when the synchronous motor is first started.

PR_1 and PR_2 are old Baldwin speaker units in which all the spring tension restraint of the armature has been removed; they thus act as very fast polar relays. An oscillographic study of PR_2 , which is essentially unloaded, shows that the micro switches are thrown to the right contacts before the end of the half cycle which actuates it. However, PR_1 , due to the loading of the shutter vane attached to the armature, is not nearly so fast, but opens in less than 3/120 second and closes in less than 4/120 second, which are the limits required for its operation. MS_1 and MS_2 are a pair of micro switches, type Z,—BZ-R, selected for near equality of spring tension. They are mounted plunger to plunger with a loose bar between them. This bar has a fulcrum at one end, and a fork at the other. Inside the fork is located the armature of PR_2 . The fork width is so adjusted that it offers no resistance to the movement of the armature except at the very end of its motion when the impact of the armature is sufficient to kick over both micro switches.

It was necessary to calibrate the neutral filters, the wedge and balancer, the diaphragm openings, and the energy at the pupil P . The filters and the wedge and balancer were measured with our photoelectric spectrophotometer (Shlaer, 1938) at the same wavelength used in the experiments, and in an analogous optical position in front of the entrance slit of the first monochromator. We first used filters and wedges made of gelatin; later they were replaced with neutral glass. The smaller diaphragms were calibrated under the microscope with a filar micrometer by measuring several diameters for each opening; the larger ones were similarly measured with a comparator.

The energy density at the pupil P was measured with a Hilger linear thermopile and a Paschen galvanometer. The thermopile was first standardized against a standard carbon filament lamp of known energy radiation. To do this we used the tube holding the pupil and the field lens, first removing the field lens and substituting

S. HECHT, S. SHLAER, AND M. H. PIRENNE

825

the slit for the pupil, and fixing the thermopile immediately behind the slit. This assembly of tube, slit, and thermopile was then mounted on an optical bench so that the standard lamp was at the specified distance of 2 meters from the receiver of the thermopile. The thermopile and its end of the tube was then covered with a thermos flask and allowed to reach thermal equilibrium. Between the source and the opening of the tube was mounted a triple leafed shutter with about 20 cm. spacing between the leaves. The surfaces facing the thermopile were blackened while those facing the source were shiny. This shutter was used to open and close the radiation to the thermopile.

The thermopile was connected to a Paschen galvanometer, which is a moving magnet type of very high sensitivity (about 2×10^{-9} volts per mm. at a meter). In series with the thermopile and galvanometer was a resistance of about 0.1 ohm, across which known potentials could be inserted to counterbalance the potential generated by the thermopile, thus using the galvanometer as a null point instrument. The radiation was first permitted to fall on the thermopile, and the galvanometer brought back to zero by means of measured counter-potentials. The radiation was then occluded and the counter-potential switched off to check the zero of the galvanometer. In this way we could measure large potentials corresponding to galvanometer swings of several meters without actually using such scale distances. The thermopile was calibrated as potential *vs.* radiant energy density incident upon its receivers for three different energy densities which covered a range of about 3 to 1, and included the actual energy density delivered by the ribbon filament lamp and the monochromators.

For calibrating the energy density through the monochromators, the field lens was replaced in the tube and the tube placed in its correct position in the apparatus. Diaphragm *D* was removed, the middle slit of the monochromator was opened to 1.5 mm., and the wedge and balancers were removed. The energy was then measured with the same thermopile and the same electrical system. With the lamp current at 19 amperes, the energy density at the pupil *P* was 27.5 microwatts per square centimeter; with the current at 18 amperes, it was 18.3 microwatts per square centimeter. In the early visual determinations we used the lamp at 19 amperes; in the later determinations at 18 amperes.

In order to convert these measurements into values of the energy at the pupil during the visual determinations, it is necessary to reduce the measured energy density by factors corresponding (*a*) to the change of the middle slit from 1.5 to 0.1 mm., (*b*) to the change in aperture of the field lens from its largest opening of 25.9 mm. diameter to the sizes of the particular diaphragms used, and (*c*) to the insertion of the wedge and balancer. All these factors were known from previous separate measurements, but we calibrated them again in their places in the apparatus by means of a sensitive dry-disc photocell in place of the thermopile behind the thermopile slit. The results merely confirmed the previous calibrations. By applying these reduction factors for the wedge at its thinnest place, the middle slit at 0.1 mm., the 10 minute diaphragm at the field lens, and the 2 mm. pupil at *P*, we found that the energy density through the pupil is 3.4×10^{-4} ergs per second when the ribbon filament lamp is running at 18 amperes. The energy calibrations were run through twice several months apart and agreed almost perfectly.

IV

Visual Measurements

From the subject's point of view, an experiment involves the report of whether or not he has seen a flash of light after he has opened the shutter for an exposure. Fixation of the red point need not be continuous, a circumstance which avoids undue fatigue. The observer is told by the operator that conditions are set and that he should try a flash when he is ready. He fixates

TABLE II
Minimum Energy for Vision

Each datum is the result of many measurements during a single experimental period, and is the energy which can be seen with 60 per cent frequency. $\lambda = 510 \text{ m}\mu$; $h\nu = 3.84 \times 10^{-12}$ ergs.

Observer	Energy	No. of quanta	Observer	Energy	No. of quanta
	<i>ergs × 10¹⁰</i>			<i>ergs × 10¹⁰</i>	
S. H.	4.83	126	C. D. H.	2.50	65
	5.18	135		2.92	76
	4.11	107		2.23	58
	3.34	87		2.23	58
	3.03	79			
	4.72	123		3.31	81
	5.68	148		4.30	112
S. S.	3.03	79	M. S.	4.61	120
	2.07	54			
	2.15	56		3.19	83
	2.38	62			
	3.69	96		3.03	79
	3.80	99		3.19	83
	3.99	104		5.30	138

the red point, and at the moment which he considers propitious, he exposes the light to his eye. The operator changes the position of the wedge, or removes or introduces a filter until he is satisfied with the precision of the measurements.

In the early measurements we considered that the threshold had been reached when the observer saw a flash of light at a given intensity six times out of ten presentations. Later the measurements were made somewhat more elaborately. Each of a series of intensities was presented many times and the frequency of seeing the flash was determined for each. From the resulting plot of frequency against intensity we chose the threshold as that amount of light which could be seen with a frequency of 60 per cent.

During 1940 and 1941 we measured the threshold for seven subjects. With

four we made several determinations each, extending over a year and a half; one subject we measured on two occasions 3 months apart; and two we measured only once. For all these observers the minimum energy necessary for vision ranges between 2.1 and 5.7×10^{-10} ergs at the cornea. These small energies represent between 54 and 148 quanta of blue-green light. The results for the individual subjects are in Table II, and are given as energy and as the number of quanta required.

It is to be noticed that these values are of the same order of magnitude as those of von Kries and Eyster, and of Barnes and Czerny, but almost twice as large. Because of the fairly wide ranges, these previous measurements and our own overlap to some extent, and it is conceivable, though not probable, that their observers may actually have needed somewhat smaller energies than ours. Chariton and Lea's results, however, are much too small. Actually their value of $17 h\nu$ is an extrapolation to zero frequency of seeing; if we take as threshold a 60 per cent frequency, their data come more nearly to $25 h\nu$. This is still too small a value, and is probably in error, as will be apparent in later sections of our paper.

V

Reflections and Absorptions

The values in Table II, as well as those of previous investigators, are the energies incident at the cornea. Nevertheless the tacit supposition has generally been made that they represent the actual energies necessary to initiate a visual act. It is important to recognize that this assumption is incorrect. Before one can know how many quanta are required to start the visual process, one must apply at least three corrections to the measurements.

The first is reflection from the cornea. This is about 4 per cent and is obviously of not much importance. The second involves loss by the ocular media between the outer surface of the cornea and the retina. It has been common opinion that this loss is small. However, the measurements of Roggenbau and Wetthauer (1927) on cattle eyes, as well as the recent measurements of Ludvigh and McCarthy (1938) on human eyes, have shown that this loss is large. From the values of Ludvigh and McCarthy it appears that at $510 m\mu$ the ocular media transmit almost exactly 50 per cent of the light entering the cornea of a young person, and less of an older one.

The next correction is much more difficult to evaluate with precision and involves the percentage of the energy absorbed by the retinal elements themselves. Since visual purple is the photosensitive substance concerned in this particular act, light which is not absorbed by it is visually useless. One cannot assume that visual purple absorbs all the light incident on the retinal cells. The fraction which it does absorb must be found by experiment.

Koenig (1894) determined the absorption of the total amount of visual

purple which can be extracted from the human eye. If this amount of visual purple is spread evenly over the whole retina, his data show that it will absorb only 4 per cent of light of 510 m μ . This is a small value. Nevertheless, it is about the same as the 4 per cent and the 13 per cent recently found by Wald (1938 b) with a similar method for the absorption of the visual purple of the rabbit and rat retinas respectively.

These figures are probably too low, first because it is unlikely that all of the visual purple in the eye has been extracted, and second, because visual purple is not evenly distributed over the retina. It is lacking in the fovea; and even in the periphery the density of the rods is known to vary in a definite way. However, these absorptions may be considered as lower limiting values.

VI

Visual Purple Absorption

We have estimated the absorption of visual purple in the retina in a completely independent manner by comparing the percentage absorption spectrum of different concentrations of visual purple with the scotopic (rod) luminosity curve of the eye measured at the retina. The comparison rests on the fact that the shape and width of the percentage absorption spectrum of a substance varies with its concentration, and that the luminosity curve must represent the percentage absorption curve of a particular concentration of visual purple in the retina.

Fig. 3 shows the absorption spectrum of frog's visual purple as determined by Chase and Haig (1938) in our laboratory, by Lythgoe (1937) in London, and by Wald (1938 b) at Harvard. The agreement of the data is obvious, and shows that the absorption spectrum of visual purple may be considered as well established. Table III gives the average of these three series of measurements computed so that the maximum density at 500 m μ has a value of 1.

From these data in Table III we may prepare a series of percentage absorption spectra for different concentrations of visual purple. Since we are not interested in the absolute concentration of visual purple, but rather in its absorption capacities, we can deal with the series of percentage absorption spectra entirely in terms of maximum absorption. It will be recalled that the photometric density d is related to the transmission I_t by the equation $d = \log(1/I_t)$, and since the absorption $I_a = 1 - I_t$, it is a simple computation to find the percentage absorption corresponding to any density value, or the reverse.

We have made such computations for a variety of visual purple densities, and Fig. 4 shows the resulting percentage absorption curves for the different maximal absorptions of visual purple. For comparisons among the curves in Fig. 4 the maxima have all been made equal to 1, but their actual values are

S. HECHT, S. SHLAER, AND M. H. PIRENNE

829

indicated in the figure. It is clear that the width of the curves increases as the concentration of visual purple increases.

The scotopic luminosity curve, as measured experimentally, records the reciprocal of the relative energy in different parts of the spectrum required for the production of a constant and very low brightness in the eye (Hecht and Williams, 1922). If this is to be compared with the absorption spectrum of

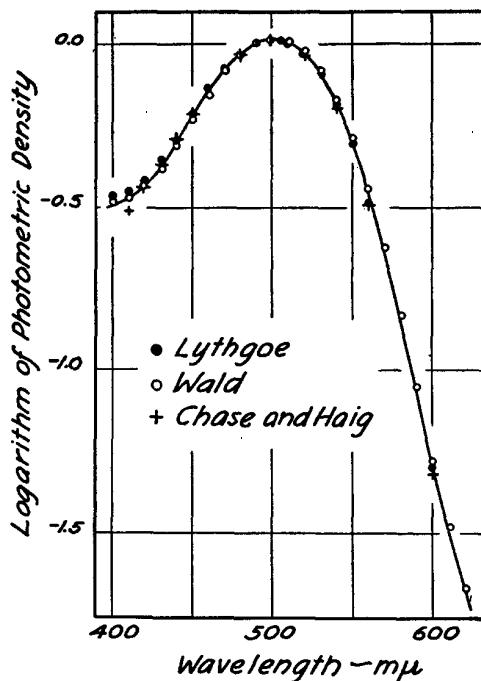


FIG. 3. Absorption spectrum of frog's visual purple. The data from the three sources have been made equal at 500 mμ.

visual purple, it must be converted into a quantum luminosity curve instead of an energy luminosity curve, because it is the number of quanta which determines the photochemical effectiveness of light and not just its energy content (Dartnall and Goodeve, 1937). Moreover, since our interest lies in retinal comparisons, the luminosity curve must be corrected for ocular media absorption in terms of the data of Ludvigh and McCarthy.

The scotopic luminosity data have been corrected in these two ways; the computed values are given in Table IV and shown as circles in Fig. 5. Included in the same figure are two percentage absorption spectra of visual purple; the

830

ENERGY, QUANTA, AND VISION

upper curve represents 20 per cent maximal absorption, while the lower curve is 5 per cent maximal absorption.

TABLE III
Absorption Spectrum of Visual Purple
 Average of data from Chase and Haig (1938), Wald (1938 b), and Lythgoe (1937).

$\lambda - m\mu$	Density	$\lambda - m\mu$	Density	$\lambda - m\mu$	Density
400	0.306	480	0.900	560	0.321
410	0.317	490	0.967	570	0.207
420	0.353	500	1.000	580	0.131
430	0.408	510	0.973	590	0.0805
440	0.485	520	0.900	600	0.0473
450	0.581	530	0.780	610	0.0269
460	0.691	540	0.628	620	0.0150
470	0.811	550	0.465		

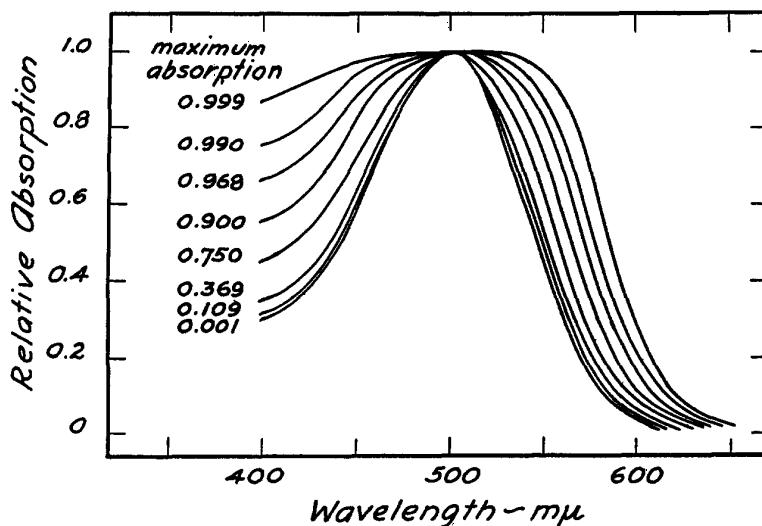


FIG. 4. Percentage absorption spectra of various concentrations of visual purple. For convenience in comparing the shapes of the curves, their maxima have all been equated to 1 and superimposed. The actual fraction absorbed at the maximum is shown for each curve. It is apparent that with increasing concentration the absorption curve steadily increases in width.

For comparing the luminosity and absorption data, it is well to confine our attention mostly to the long wave half of the luminosity curve because of the larger number of points involved. From the comparison it is apparent that the 5 per cent maximum absorption curve describes the points quite well, but

S. HECHT, S. SHLAER, AND M. H. PIRENNE

831

TABLE IV
Rod Luminosity Distribution in Spectrum

The original energy luminosity data of Hecht and Williams (1922) in column 2, when divided by the corresponding wavelengths in column 1, yield the quantum luminosity values in column 3 after being multiplied by a factor so that the maximum at 511 m μ equals 1. When these values in column 3 are divided by the ocular media transmission data in column 4 from Ludvigh and McCarthy (1938), they yield the spectral luminosity distribution at the retina given in column 5 after multiplication by a factor so that the maximum at 502 m μ is 1.

$\lambda - m\mu$	Energy luminosity at cornea	Quantum luminosity at cornea	Ocular transmission	Quantum luminosity at retina
412	0.0632	0.0779	0.116	0.336
455	0.399	0.447	0.410	0.545
486	0.834	0.874	0.472	0.926
496	0.939	0.964	0.490	0.984
507	0.993	0.998	0.506	0.986
518	0.973	0.957	0.519	0.921
529	0.911	0.877	0.540	0.812
540	0.788	0.743	0.559	0.665
550	0.556	0.515	0.566	0.455
582	0.178	0.155	0.596	0.131
613	0.0272	0.0226	0.625	0.0181
666	0.00181	0.00139	0.672	0.00104

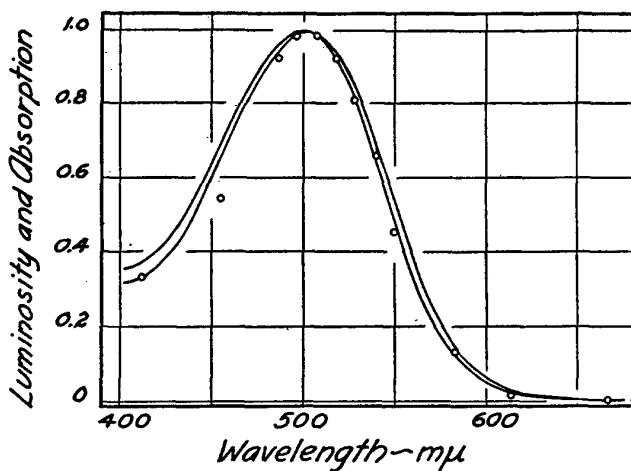


FIG. 5. Comparison of scotopic luminosity at the retina with visual purple absorption. The points are the data of Hecht and Williams corrected for quantum effectiveness and ocular media transmission. The curves are the percentage absorption spectra of visual purple; the upper curve represents 20 per cent maximal absorption, and the lower one 5 per cent maximal absorption. All curves have been made equal to 1 at the maximum, 500 m μ , for ease in comparison.

that the 20 per cent curve is definitely excluded, because its absorption on both sides is just too high. The 10 per cent absorption curve, not shown in the figure, is perhaps slightly better than the 5 per cent one; it cuts through more points. In any case, both values are of the same order of magnitude as those found by Koenig and by Wald. However, to be quite safe, we may take 20 per cent as the upper limit for the absorption of $510 \text{ m}\mu$ by the visual purple in the human retina after complete dark adaptation.

VII

Energy Absorbed by the Rods

It is clear now why the 54 to 148 quanta required at the cornea cannot represent the energy actually employed in vision. About 4 per cent of this incident light is reflected by the cornea; almost precisely 50 per cent is absorbed by the lens and other ocular media; and of the rest, at least 80 per cent passes through the retina without being absorbed. If corrections are made for these factors, the range of 54 to 148 quanta at the cornea becomes as an upper limit 5 to 14 quanta absorbed by the visual purple of the retina.

Visual purple is in the terminal segments of the rods, and the 10 minute circular visual field contains about 500 rods (Østerberg, 1935). Since the number of absorbed quanta is so small, it is very unlikely that any one rod will take up more than one quantum. In fact, the simplest statistical considerations show that if 7 quanta are absorbed by 500 rods, there is only a 4 per cent probability that 2 quanta will be taken up by a single rod. We may therefore conclude that in order for us to see, it is necessary for only 1 quantum of light to be absorbed by each of 5 to 14 retinal rods.¹

It is very likely that the photodecomposition of visual purple in solution has a quantum efficiency of 1 (Dartnall, Goodeve, and Lythgoe, 1938). Our data then mean that 1 molecule of visual purple needs to be changed simultaneously in each of 5 to 14 rods, in order to produce a visual effect. This is indeed a small number of chemical events, but by virtue of its very smallness, its reality may be tested in an entirely independent manner.

VIII

Poisson Distributions

The energy calibration of the light gives merely the average number of quanta per flash. This is in the nature of the measurement, because the

¹ These data disprove the supposition made by Granit, Holmberg, and Zewi (1938) that most of the visual purple in the retina is inert as sensory substance, and that sensory impulses from the rods are "initiated by the bleaching of a thin surface film, which had to contain only an immeasurably small fraction of the total quantity present" (Granit, Munsterhjelm, and Zewi, 1939). Since the maximum visual purple concentration which the retina can achieve is able to absorb only 5 to 14 quanta at the threshold of vision, a very small fraction of the total visual purple would absorb much less than one quantum and would be ineffective for visual purposes.

S. HECHT, S. SHLAER, AND M. H. PIRENNE

833

thermopile records only the energy density, which is the number of quanta per second from a continuously incident light. Each flash, however, will not always deliver this average number. Sometimes the flash will yield fewer, sometimes more, quanta.

Since absorption of this group of quanta by the retina represents discrete and independent events which occur individually and collectively at random, the actual number of such retinal events which any given flash provides will vary according to a Poisson probability distribution (Fry, 1928). Let n be the number of quanta which it is necessary for the retina to absorb in order for us

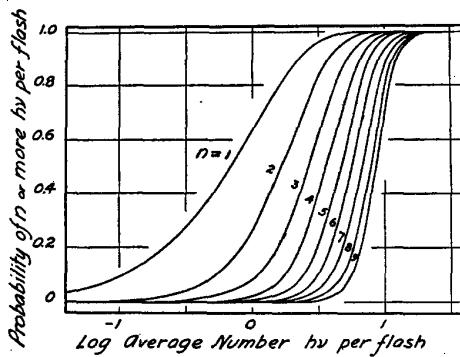


FIG. 6. Poisson probability distributions. For any average number of quanta ($h\nu$) per flash, the ordinates give the probabilities that the flash will deliver to the retina n or more quanta, depending on the value assumed for n .

to see a flash of light. Let a be the average number of quanta which any flash yields to the retina. Then the Poisson distribution states that

$$P_n = a^n / e^a n!$$

in which P_n is the probability that the flash will yield the necessary n quanta, and e is the base of natural logarithms. A special virtue of the Poisson distribution is that it has only one parameter, and is thus determined when the average number a is set. The values of P_n for various values of a and n are available in printed tables (e.g. Fry, 1928).

Since for us to see a flash of light the retina must absorb n quanta, we shall also see when the retina absorbs more than n quanta. From the published Poisson distributions, one can then compute the probability that n or more quanta will be delivered to the retina in a given flash when the average number of quanta delivered by that flash is known. The values computed in this way for different values of a and n are shown in Fig. 6.

There are two significant features of Fig. 6. One is that the shape of the distributions is fixed and different for every value of n . The curve becomes

steeper as n increases. It follows from this that if the probability distribution could be determined by experiment, its shape would automatically reveal the value of n corresponding to it.

Another and equally important feature of Fig. 6 is that the relationship is expressed in terms of the logarithm of the average number of quanta per flash. Therefore, for comparison with the distributions in Fig. 6, the experiments need not employ the absolute values of the average number of quanta delivered per flash, but merely their relative values.

The experiments may then be made quite simply. On many repetitions of a flash of given average energy content, the frequency with which the flash is seen will depend on the probability with which it yields n or more quanta to the retina. When this frequency is measured for each of several intensities, a distribution is secured whose shape, when plotted against the logarithm of the average energy content, should correspond to one of the probability distributions in Fig. 6, and should thus show what the value of n has been.

IX

Frequency of Seeing

We have made determinations of this kind. The experimenter varies the intensity of the light by placing the wedge in specific positions unknown to the observer. The observer then elicits the flash whenever he is ready, and merely reports whether he has seen it or not. The intensities are presented in a deliberately random sequence, each for a specific number of times, usually 50. The procedure is simplified for the operator by a series of accurately made stops against which the wedge may be rapidly set in predetermined positions. A complete series in which six intensities are used requires about $1\frac{1}{2}$ hours of continuous experimentation composed of two or three periods of intensive work.

The comfort of the observer is of great importance and this must be at a maximum. It is equally important that fixation should not be rigidly continuous because this is fatiguing. Above all, the observer must be on guard to record any subjective feelings of fatigue the moment they become apparent. The experiment is much facilitated by the fact that the observer controls the occurrence of the flash, and can set it off only when he is thoroughly fixated and ready for an observation.

The data for the three observers who engaged in this experiment are given in Table V. One experiment for each observer is plotted in Fig. 7. The points in the figure record the percentage frequency with which a flash of light is seen for flashes of average quantum content shown in the abscissas. Comparison with the curves in Fig. 6 shows that the measurements are best fitted by Poisson distributions in which n is 5, 6, and 7 quanta per flash. For the two other

S. HECHT, S. SHLAER, AND M. H. PIRENNE

835

experiments in Table IV, n is 7 and 8. No special statistical methods are necessary to determine which curve fits the data, since smaller and larger values of n are easily excluded by the simplest visual comparison.

TABLE V
Energy and Frequency of Seeing

Relation between the average number of quanta per flash at the cornea and the frequency with which the flash is seen. Each frequency represents 50 flashes, except for S. H. for whom there were 35 and 40 for the first and second series respectively.

S. H.		S. H.		S. S.		S. S.		M. H. P.	
No. of quanta	Frequency								
<i>per cent</i>		<i>per cent</i>		<i>per cent</i>		<i>per cent</i>		<i>per cent</i>	
46.9	0.0	37.1	0.0	24.1	0.0	23.5	0.0	37.6	6.0
73.1	9.4	58.5	7.5	37.6	4.0	37.1	0.0	58.6	6.0
113.8	33.3	92.9	40.0	58.6	18.0	58.5	12.0	91.0	24.0
177.4	73.5	148.6	80.0	91.0	54.0	92.9	44.0	141.9	66.0
276.1	100.0	239.3	97.5	141.9	94.0	148.6	94.0	221.3	88.0
421.7	100.0	386.4	100.0	221.3	100.0	239.3	100.0	342.8	100.0

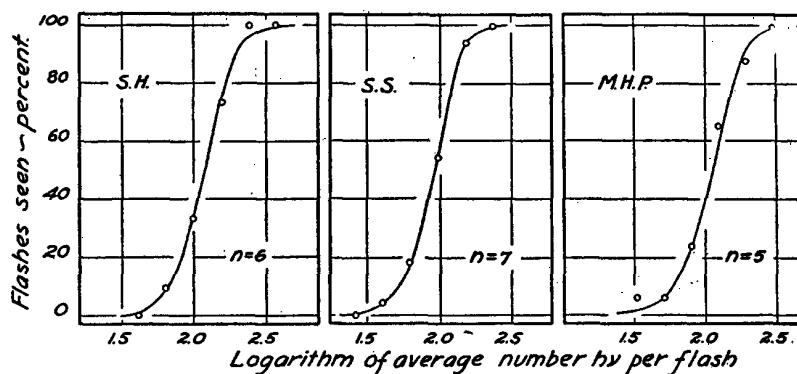


FIG. 7. Relation between the average energy content of a flash of light (in number of $h\nu$) and the frequency with which it is seen by three observers. Each point represents 50 flashes, except for S.H. where the number is 35. The curves are the Poisson distributions of Fig. 6 for n values of 5, 6, and 7.

From these measurements it is apparent that the number of critical events in the retina required to produce a visual effect lies between 5 and 8. These values are in such good agreement with the results determined by the straightforward physical measurements already described that we must consider them as the actual number of quanta absorbed by the retina.

X

Physical Fluctuation and Biological Variation

It is unimportant that the number of quanta delivered to the cornea is very much higher than the number finally involved in vision according to these measurements. This is because most of the light incident on the cornea is wasted and does not contribute to the initiation of a visual act. The amount falling on the cornea could be greatly increased by any arrangement in the eye which would act as a filter. Thus, the cornea and the lens might be pigmented, and this probably contributes to the fact that the oldest investigator (S.H.) actually requires the highest number of quanta incident on the cornea. Indeed, one might even put a filter immediately in front of the eye since the precise position of the filter in the optical system is immaterial. Nevertheless, the probability distributions would still remain the same, and by their shape would yield the magnitude of the number of events involved in the visual act.

It is necessary to amplify this point somewhat. Fluctuations are part of all physical systems, but they become significantly large only when the number of individual events, in the modern physical sense, is small. The general phenomenon is known as the shot-effect and has been studied extensively in electron emission, though it has wide application in the problem of measurements (Schottky, 1922; Barnes and Czerny, 1932). As a rough approximation, one may say that the range of variation is proportional to the square root of the number of individual events involved in the process.

In the optical system of our apparatus, the light from the ribbon filament lamp varies in intensity from moment to moment, but because the number of quanta emitted is enormous, the variation is almost too small to be measured. However, when the light intensity has been reduced first by the filters and wedge, then by the monochromators, then by the shutter, then by the ocular media, and finally by the retina itself, it has become so low that it represents only a few quanta per flash, and is therefore subject to great variation.

Barnes and Czerny (1932), and following them Brumberg and Vavilov (1933) realized that fluctuations must occur in the energy necessary for vision, and both groups of investigators looked for them. But they both missed the point of where the source of the fluctuations is and supposed it to be the energy deposited at the cornea. Brumberg and Vavilov even expected differences in the fluctuations for different wavelengths because of the greater energy required for seeing red light, for example, than blue-green light in conformity with the scotopic visibility curve of Fig. 5. However, the comparisons in Fig. 5 show that the differences in number of quanta required for vision in different parts of the spectrum record merely their relative absorption by visual purple. The number of absorbed quanta for an ultimate effect is the same regardless of wavelength and it is this number which sets the magnitude of the physical fluctuation encountered.

S. HECHT, S. SHLAER, AND M. H. PIRENNE

837

In deriving the curves of Fig. 6 for the quantitative statement of this physical fluctuation in terms of the Poisson probability distribution, we have made the single assumption that a constant number of quanta n must be absorbed by the retina in order for us to see a flash of light. Since it is conceivable, in view of the variability of an organism from moment to moment, that this value n is not constant, we have considered the consequences of assuming that the number n varies from time to time. The results show that biological variation is a factor of no great importance.

The situation may be best made clear by an example. Suppose that instead of n being constant, it varies between 4 and 8 quanta per visual act, and that the frequency with which 4, 5, 6, 7, and 8 quanta are necessary is distributed in terms of an ordinary probability distribution. The curves in Fig. 6 representing the frequency distributions for various values of n may then be weighted in this way and averaged. The average curve which is then secured is practically the same as the original Poisson distributions in Fig. 6, and may be fitted by the curves for $n = 4$ or 5.

Thus, when biological variation is imposed upon the physical variation, there is no change in the essential characteristics of the physical distribution. Instead, the value of n merely falls below the average of the biological distribution, and is never below the lowest value in the distribution. This tells us that when, as in Fig. 7, the measurements yield n values of 5, 6, or 7, these numbers represent lower limiting values for the physical number of quanta. In other words, the only effect which biological variation has on the physical variation is to decrease the slope of the curves in Fig. 7 and thus make the apparent number of quanta smaller than the real number.

These considerations serve for understanding the meaning of the fluctuations shown by an organism in its response to a stimulus. It has generally been assumed that a constant stimulus, when presented frequently, remains constant, and that the fluctuations in response are an expression of the variations undergone by the organism. Indeed, this is one of the tenets of psychological measurements, and an elaborate structure of psychometrics has grown up on it as a basis (*cf.* Guilford, 1936).

The present evaluation of our measurements shows, however, that at the threshold the emphasis has been in the wrong place. At the threshold where only a few quanta of energy are involved, it is the stimulus which is variable, and the very nature of this physical variability determines the variation encountered between response and stimulus. Moreover, even when biological variation is introduced, it is the physical variation which essentially dominates the relationship.

This is at the absolute threshold. One may wonder, however, whether a differential threshold at any level of intensity may also involve a small number of events which determines the differentiation, and which may therefore be

subject to a similar physical variation as at the absolute threshold itself. Only experiment can decide this.

The fact that for the absolute visual threshold the number of quanta is small makes one realize the limitation set on vision by the quantum structure of light. Obviously the amount of energy required to stimulate any eye must be large enough to supply at least one quantum to the photosensitive material. No eye need be so sensitive as this. But it is a tribute to the excellence of natural selection that our own eye comes so remarkably close to the lowest limit.

SUMMARY

1. Direct measurements of the minimum energy required for threshold vision under optimal physiological conditions yield values between 2.1 and 5.7×10^{-10} ergs at the cornea, which correspond to between 54 and 148 quanta of blue-green light.

2. These values are at the cornea. To yield physiologically significant data they must be corrected for corneal reflection, which is 4 per cent; for ocular media absorption, which is almost precisely 50 per cent; and for retinal transmission, which is at least 80 per cent. Retinal transmission is derived from previous direct measurements and from new comparisons between the percentage absorption spectrum of visual purple with the dim-vision luminosity function. With these three corrections, the range of 54 to 148 quanta at the cornea becomes as an upper limit 5 to 14 quanta actually absorbed by the retinal rods.

3. This small number of quanta, in comparison with the large number of rods (500) involved, precludes any significant two quantum absorptions per rod, and means that in order to produce a visual effect, one quantum must be absorbed by each of 5 to 14 rods in the retina.

4. Because this number of individual events is so small, it may be derived from an independent statistical study of the relation between the intensity of a light flash and the frequency with which it is seen. Such experiments give values of 5 to 8 for the number of critical events involved at the threshold of vision. Biological variation does not alter these numbers essentially, and the agreement between the values measured directly and those derived from statistical considerations is therefore significant.

5. The results clarify the nature of the fluctuations shown by an organism in response to a stimulus. The general assumption has been that the stimulus is constant and the organism variable. The present considerations show, however, that at the threshold it is the stimulus which is variable, and that the properties of its variation determine the fluctuations found between response and stimulus.

BIBLIOGRAPHY

- Ångström, K., Energy in the visible spectrum of the Hefner standard, *Physic. Rev.*, 1903, 17, 302.

- Barnes, R. B., and Czerny, M., Lässt sich ein Schroteffekt der Photonen mit dem Auge beobachten?, *Z. Physik.*, 1932, **79**, 436.
- Brumberg, E., and Vavilov, S., Visuelle Messungen der statistischen Photonenschwankungen, *Bull. Acad. Sc. U.R.S.S.*, 1933, 919.
- Buisson, H., The minimum radiation visually perceptible, *Astrophys. J.*, 1917, **46**, 296.
- Chariton, J., and Lea, C. A., Some experiments concerning the counting of scintillations produced by alpha particles. Part I, *Proc. Roy. Soc. London, Series A*, 1929, **122**, 304.
- Chase, A. M., and Haig, C., The absorption spectrum of visual purple, *J. Gen. Physiol.*, 1938, **21**, 411.
- Dartnall, H. J. A., and Goodeve, C. F., Scotopic luminosity curve and the absorption spectrum of visual purple, *Nature*, 1937, **139**, 409.
- Dartnall, H. J. A., Goodeve, C. F., and Lythgoe, R. J., The effect of temperature on the photochemical bleaching of visual purple solutions, *Proc. Roy. Soc. London, Series A*, 1938, **164**, 216.
- du Notiy, P. Lecomte, Energy and vision, *J. Gen. Physiol.*, 1921, **3**, 743.
- Fry, T. C., Probability and its engineering uses, New York, Van Nostrand, 1928, 476.
- Graham, C. H., and Margaria, R., Area and the intensity-time relation in the peripheral retina, *Am. J. Physiol.*, 1935, **113**, 299.
- Granit, R., Holmberg, T., and Zewi, M., On the mode of action of visual purple on the rod cell, *J. Physiol.*, 1938, **94**, 430.
- Granit, R., Munsterhjelm, A., and Zewi, M., The relation between concentration of visual purple and retinal sensitivity to light during dark adaptation, *J. Physiol.*, 1939, **96**, 31.
- Grijns, G., and Noyons, A. K., Ueber die absolute Empfindlichkeit des Auges für Licht, *Arch. Anat. u. Physiol. Abt.*, 1905, 25.
- Guilford, J. P., Psychometric methods, New York, McGraw-Hill, 1936.
- Hecht, S., The quantum relations of vision, *J. Opt. Soc. America*, 1942, **32**, 42.
- Hecht, S., Haig, C., and Chase, A. M., The influence of light adaptation on subsequent dark adaptation of the eye, *J. Gen. Physiol.*, 1937, **20**, 831.
- Hecht, S., Shlaer, S., and Pirenne, M. H., Energy at the threshold of vision, *Science*, 1941, **93**, 585.
- Hecht, S., and Williams, R. E., The visibility of monochromatic radiation and the absorption spectrum of visual purple, *J. Gen. Physiol.*, 1922, **5**, 1.
- Ives, H. E., The minimum radiation visually perceptible, *Astrophys. J.*, 1916, **44**, 124.
- Koenig, A., Ueber den menschlichen Sehpurpur und seine Bedeutung für das Sehen, *Sitzungsber. k. Akad. Wissenschaft.*, Berlin, 1894, 577.
- von Kries, J., and Eyster, J. A. E., Über die zur Erregung des Sehorgans erforderlichen Energiemenzen, *Z. Sinnesphysiol.*, 1907, **41**, 394.
- Langley, S. P., The bolometer and radiant energy, *Proc. Am. Acad. Sc.*, 1881, **16**, 342.
- Langley, S. P., Energy and vision, *Phil. Mag.*, 1889, **27**, series 5, 1.
- Ludvigh, E., and McCarthy, E. F., Absorption of visible light by the refractive media of the human eye, *Arch. Ophth.*, Chicago, 1938, **20**, 37.
- Lythgoe, R. J., The absorption spectra of visual purple and of indicator yellow, *J. Physiol.*, 1937, **89**, 331.
- Østerberg, G., Topography of the layer of rods and cones in the human retina, *Acta Ophth., Copenhagen*, 1935, suppl. 6, 106 pp.

840

ENERGY, QUANTA, AND VISION

- Piper, H., Über Dunkeladaptation, *Z. Psychol. u. Physiol. Sinnesorgane*, 1903, **31**, 161.
- Reeves, P., The minimum radiation visually perceptible, *Astrophys. J.*, 1917, **46**, 167.
- Rogggenbau, C., and Wetthauer, A., Über die Durchlässigkeit der brechenden Augenmedien für langwelliges Licht nach Untersuchungen am Rindsauge, *Klin. Monatsbl. Augenheilk.*, 1927, **79**, 456.
- Russell, H. N., The minimum radiation visually perceptible, *Astrophys. J.*, 1917, **45**, 60.
- Shottky, W., Zur Berechnung und Beurteilung des Schroteffektes, *Ann. Physik.*, 1922, **68**, 157.
- Shlaer, S., A photoelectric transmission spectrophotometer for the measurement of photosensitive solutions, *J. Opt. Soc. America*, 1938, **28**, 18.
- Wald, G., Area and visual threshold, *J. Gen. Physiol.*, 1938 *a*, **21**, 269.
- Wald, G., On rhodopsin in solution, *J. Gen. Physiol.*, 1938 *b*, **21**, 795.
- Wentworth, H. A., A quantitative study of achromatic and chromatic sensitivity from center to periphery of the visual field, *Psychological Monographs*, No. 183, Princeton, New Jersey, and Albany, New York, *Psychological Review Co.*, 1930, **40**, 189 pp.
- Zwaardemaker, H., Die physiologisch wahrebmbaren Energiewanderungen, *Ergebn. Physiol.*, 1905, **4**, 423.

Chapter 18

Photon Correlation Measurements with Beamsplitter and Photomultipliers (Lajos Jánossy)



Sándor Varró

Abstract This chapter includes a brief biography of Lajos Jánossy (1912–1978) presented by Sándor Varró (Sect. 18.1) and his translation into English from Hungarian of Jánossy's group paper of 1954 year “Coincidences of photons traveling in coherent beams of light” (Sect. 18.2). Starting with cosmic rays research and the study of statistics of coincidences in Geiger counters, Jánossy brought some of these techniques to investigation of photon correlations using a beamsplitter and photomultipliers. Although no correlations were observed in these experiments, Jánossy's papers of 1954–1955 are considered as forerunners of the Hanbury Brown–Twiss type correlations with visible light.

18.1 Lajos Jánossy: Biography

Born in Budapest in 1912, since the age of 6, Lajos Jánossy lived abroad moving with his parents: he attended university in Vienna (1930–1934) and, later, the Humboldt University in Berlin. He worked in the laboratory of Werner Kolhörster in Berlin (1934–1936), in the field of astrophysics. In 1936, fleeing Nazism, he moved to London, until 1938, carrying out research at Birkbeck College. From 1938 he worked at the University of Manchester under the subsequent Nobel laureate Patrick Blackett on astrophysics, heading the cosmic radiation research group. In 1940 and 1941, he discovered the penetrative showers of cosmic radiation and identified their most important properties and their relationship to the extended particle showers present in the air. Based on this, the view was established that, under the effect of the primarily high-energy cosmic particles, an extremely large number of particles—primarily mesons—with high penetration capability are created in the upper layers of the atmosphere. As a result of repeated interactions,

S. Varró (✉)

Department of Theoretical Solid State Physics, Institute for Solid State Physics and Optics, Wigner Research Centre for Physics, Hungarian Academy of Sciences, Konkoly-Thege Miklós út 29-33, Budapest 1121, Hungary
e-mail: varro.sandor@wigner.mta.hu

a very large number of particles (showers) reach the surface of the earth. His book *Cosmic Rays* (The Clarendon Press, Oxford), 1950 (second edition) is the result of this work (Fig. 18.1).

During 1944, Erwin Schrödinger and Jánossy exchanged several letters (see books [1, 2] (Jánossy attended earlier some of Schrödinger's courses in Berlin [1])). Schrödinger had become interested in a note by Jánossy in Nature on the statistics of coincidences in Geiger counters. Jánossy was invited to lecture at a Dublin summer school on cosmic rays in 1945, and in 1947 he was invited by Walter Heitler and Erwin Schrödinger to join the Dublin Institute for Advanced Studies as a senior professor and group leader of the cosmic rays research laboratory for nearly three years (Fig. 18.2).

Invited later by the Hungarian government, Jánossy returned home in 1950, not only for reasons of promised scientific possibilities, but also because his foster-father (György Lukács philosopher) and his mother had also returned home from their emigration in Moscow. In Budapest Jánossy was charged with the task of managing the Cosmic Radiation Department at the Central Research Institute for Physics (Hungarian abbreviation: KFKI) founded in 1950. He was very active in scientific organization, education and public life. He was appointed deputy director of the KFKI from 1950 to 1956, and the director from 1956 to 1970. In addition, he was also active in university education: the Department of Nuclear Physics at Loránd Eötvös University was established for him. He was the first head of this Department from 1957 to 1970. Jánossy was elected member (1950) and, later, vice-president (1961–1973) of the Hungarian Academy of Sciences, member of the Bulgarian Academy of Sciences, the Irish Academy of Sciences, the Mongolian

Fig. 18.1 Lajos Jánossy
(from BME OMIKK
Historical Archive of Science
and Technology, Hungary)





Fig. 18.2 Jánossy, Born and Schrödinger (with permission of the Institute for Solid State Physics and Optics, Hungarian Academy of Sciences)

Academy of Sciences and the Academy of Sciences of the German Democratic Republic.

Until the fifties, the most important field in the research of high-energy particles was the examination of cosmic radiation. But as the large accelerators started to take over the leading role, Jánossy turned away from the examination of cosmic radiation and took the theoretical problems of quantum mechanics, the dual character of light as well as the theory of relativity in hand.

Jánossy and co-workers measured the photon correlation by using (for the first time) photomultipliers and the photon-counting technique [3–7]. Although no correlations were observed, this part of their study is considered as a forerunner of the Hanbury Brown–Twiss type correlations with visible light. Jánossy and co-workers also studied single photons interference in highly attenuated beam [7–10]. In relation to one of Jánossy photon experiments, Schrödinger was not slow to write a letter in which he emphasized the importance of this work. Jánossy chapter of the book [7] gives a retrospective of his experiments with highly attenuated beams.

In the last one and a half decades of his theoretical activity, Jánossy was engaged in the hydrodynamic model of quantum mechanics and the interpretation problems of the theory of relativity. He educated students to the theory of relativity and wrote the book *Theory of relativity based on physical reality* (Akadémiai Kiadó, 1971) with his own interpretation of the theory based on Lorentz transformation that is different from Einstein's interpretation accepted by a wide scientific community.

References

1. W. Moore, *Schrödinger: Life and Thought*, p. 416, Cambridge University Press (2015).
2. *In memoriam Lajos Jánossy 75, Erwin Schrödinger 100*, Budapest: MTA Központi Fizikai Kutató Intézete: Ungarische Akademie der Wissenschaften, Zentralforschungsinstitut für Physik, 1987.
3. A. Ádám, L. Jánossy and P. Varga: Coincidences of photons traveling in coherent beams of light (In Hungarian; Koherens fénynyalábokban haladó fotonok koincidencái), *Magyar Fizikai Folyóirat* **2** 499–510 (1954).
4. A. Ádám, L. Jánossy and P. Varga: Beobachtungen mit dem Elektronenvervielfacher an kohärenten Lichtstrahlen. *Annalen der Physik* **16** 408–413 (1955).
5. A. Ádám, L. Jánossy and P. Varga: Coincidences between photons contained in coherent light rays. (in Russian), *Acta Physica Hungarica* **4**, 301–305 (1955).
6. G. Farkas, L. Jánossy, Z. Náray and P. Varga: Intensity correlation of coherent light beams, *Acta Physica Hungarica* **18**, 199–212 (1964).
7. L. Jánossy: Experiments and theoretical consideration concerning the dual nature of light, in *Cooperative Phenomena*, H. Haken and M. Wagner, Eds, Springer-Verlag, Berlin (1973).
8. L. Jánossy and Z. Náray: Investigations on the dual nature of light (In Hungarian: A fény kettős természetére vonatkozó vizsgálatok), *Fizikai Szemle* **8**, 3–9 (1958).
9. L. Jánossy and Z. Náray: The interference phenomena of light at very low light intensities, *Acta Physica Hungarica* **7** 403–424 (1957).
10. L. Jánossy and Z. Náray: Investigations into interference phenomena at extremely low light intensities by means of a large Michelson interferometer, *Suppl. Nuovo Cimento* **9**, 369–384 (1958).

18.2 Coincidences of Photons Traveling in Coherent Beams of Light: Translation of an Original Paper by A. Ádám, L. Jánossy and P. Varga

Translation from Hungarian by Sándor Varró from Ádám András, Jánossy Lajos and Varga Péter: Koherens fénynyalábokban haladó fotonok koincidencái. *Magyar Fizikai Folyóirat* **2**, 499–510 (1954). “Manuscript received on 17 August 1954.”

This reference in English: A. Ádám, L. Jánossy and P. Varga: Coincidences of photons traveling in coherent beams of light. *Hungarian Physical Journal* **2**, 499–510 (1954) (in Hungarian).

See also paper in German: A. Ádám, L. Jánossy and P. Varga: Beobachtungen mit dem Elektronenvervielfacher an kohärenten Lichtstrahlen. *Annalen der Physik* **16** 408–413 (1955).

18.2.1 Introduction

The aim of our investigation is to answer the question whether photons traveling in coherent beams of light are independent or not? The quantum theory states that such photons are independent in the following sense: if we let a ray of light to fall on a partly transmitting and partly reflecting mirror, then the ray of light is divided into one transmitted and one reflected rays, and by letting the two rays impinge on photon counters, only one or the other photon counter will be activated. The [experimental] arrangement with which we shall investigate this statement is the following (Fig. 18.3).

From the light source F the light falls on a partly transmitting and partly reflecting mirror T . The light ray 1 reflected by the mirror falls to the photomultiplier M_1 , and the transmitted ray 2 falls on the photomultiplier M_2 . We observe the coincidences of the two multipliers by a suitable coincidence apparatus.

The connections of the coincidence experiments with quantum mechanics and with the wave-particle problem, respectively, has been treated in details by Jánossy [1]. In this same paper the author has pointed to the importance of the experimental investigation of this question.

18.2.2 Accidental Coincidences and the Choice of the Coincidence Resolution Capability

In the above arrangement we receive coincidences, even if the coherent light rays do not cause systematic coincidences. Since, when two bursts arrive within a certain time τ (where τ is the resolution time of the coincidence apparatus), then the apparatus counts, which means that there are accidental coincidences generated.

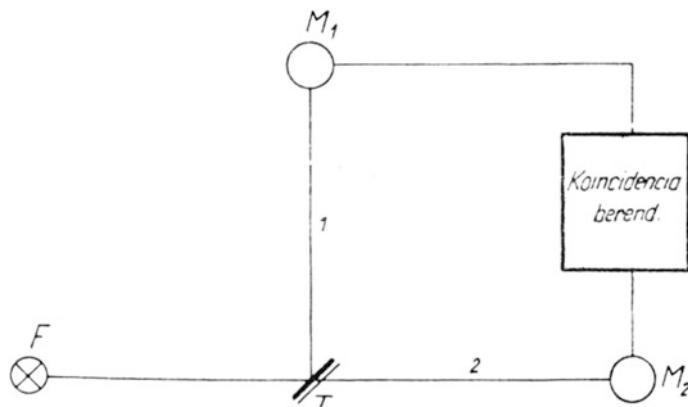


Fig. 18.3 Schematic arrangement of the coincidence experiment

In our experiment we have to investigate whether, in the case of applying coherent light rays, more coincidences appear between the two photomultipliers, than accidental coincidences.

Primarily, we have to prove that, in spite of the small probability of the activation of the photomultiplier, it is possible to surely prove the difference between the systematic and accidental coincidences.

If the quantum efficiency of the photomultiplier is p , the average intensity is n photons/s, and the number of counts initiated by the dark current is N_s /s, then

$$N = pn + N_s \quad (18.1)$$

will be the number of pulses per seconds, appearing on the anode of the photomultiplier. In our measurements we have cooled down the photomultipliers to the temperature of liquid nitrogen. In this way we have managed that the number of counts stemming from the dark current was negligible to the number of counts caused by the light. In this case $N = pn$. As is known, in a measurement during time duration t , the number of accidental coincidences is given by

$$K_v = 2N^2\tau \cdot t, \quad (18.2)$$

where τ is the resolution capability of the coincidence apparatus. (For simplicity, we have assumed that both the quantum efficiencies of the two photomultipliers and the light intensities on them are the same. The calculation is similar if we assume different quantum efficiencies and intensities.)

The number of all observed coincidences are equal to the sum of the accidental coincidences and systematic coincidences. In the determination of the number of systematic coincidences one has to distinguish two cases. On one hand, a photon may not activate both of the photomultipliers because in one of the photomultipliers accidentally no electron emission occurs (the probability of that in both photomultiplier emission occurs is p^2). On the other hand, the photon cannot at all activate both photomultipliers, because it is traveling exclusively along one or the other beam.

If we assume that an ε is a part of the photons, that is $\varepsilon \cdot n$ photons are able to create systematic coincidences, then the number of the systematic coincidences are $\varepsilon \cdot p^2 n \cdot t$. Then the number of all coincidences are

$$K = K_v + \varepsilon \cdot p^2 n \cdot t = 2N^2\tau \cdot t + \varepsilon \cdot pN \cdot t. \quad (18.3)$$

If most of the photons would create systematic coincidences, i.e., when $\varepsilon \approx 1$ would hold, then, during a relatively short measurement time we could check whether there are systematic coincidences or not. The deviation between the number of coincidences may be said to be significant if

$$K > K_v + 3\Delta K_v, \quad (18.4)$$

(where ΔK_v is the error of the number of accidental coincidences). This is because the probability of occurrence of fluctuations larger than three times of the error ΔK_v is negligibly small. If we insert to expression (18.4) the error expected on the basis of Poisson distribution, we have $(\Delta K_v)^2 = K_v$, and now we know for how long should the measurement be performed.

$$\varepsilon \cdot pN \cdot t > 3\sqrt{2N^2\tau \cdot t}$$

and

$$t > \frac{18}{\varepsilon^2 p^2} \cdot \tau. \quad (18.5)$$

With a usual coincidence arrangement one can easily manage to have a resolution time $\tau = 2 \times 10^{-6}$ s. When we insert this value into (18.5), by taking $p = 3 \times 10^{-3}$ [2] and $\varepsilon = 1$, then we obtain $t > 4$ s.

From the previous considerations it can be seen, that if most of the photons created systematic coincidences, then this could be demonstrated by measurements within few seconds.

If only a small part of the photons create systematic coincidences, i.e. if $\varepsilon \ll 1$, then for the proof of the occurrence of systematic coincidences, the measurement time should be relatively long. The question is whether in case of a realistic measurement time $[T]$ what kind of statements can we make? This is determined by the error of ε . From (18.3) we have

$$\varepsilon = \frac{K - K_v}{pN \cdot T}, \quad (18.6)$$

and

$$\Delta\varepsilon = \pm \frac{1}{pN \cdot T} \Delta(K - K_v) = \pm \frac{\sqrt{2K_v}}{pN \cdot T} = \pm \frac{2}{p} \sqrt{\frac{\tau}{T}}. \quad (18.7)$$

In deriving the above two expressions, on the basis of the Poisson distribution, we have replaced the errors by the square roots of the averages, and because of the assumption $\varepsilon \ll 1$, we have used the approximation $K \approx K_v$. By putting the former values for the parameters into (18.7), and assuming measurement time T of the order of hours, we have $\Delta\varepsilon = \pm 7 \times 10^{-3}$. This means that with a measurement time $T = 5$ h it will be possible to find whether more than 2% of the photons generate systematic coincidences.

It is seen that with a relatively small quantum efficiency of the photomultipliers, and with a usual coincidence apparatus temporal resolution ($\tau = 2 \times 10^{-6}$ s), it is possible to probe whether there are more coincidences as the accidental ones, or there are also systematic coincidences, respectively.

18.2.3 The Method of the Measurement

In order that we could decide, whether there are systematic coincidences appearing, we have to know the accurate number of the accidental coincidences. The measurement should be performed in such a way that we directly measure the accidental coincidences. When we shine the photomultipliers with two separate light sources, then only accidental coincidences will appear. It is purposeful to perform the measurement by using alternating coherent and incoherent excitations, and directly compare the obtained number of coincidences. (See Fig. 18.4; F and F_1 , F_2 are shining alternatively).

18.2.4 The Optical Arrangement

The optical arrangement used by the experiment can be seen in Fig. 18.5.

By using a prism monochromator, from the spectrum of the light source F we have chosen a range around 5000 Å. We have let the outgoing light through the lens L fall on the mirror T , which directs this light to photomultipliers M_1 , M_2 . The light from the light sources F_1 and F_2 travels to M_1 and M_2 , respectively, immediately besides the mirror T . In front of the photomultipliers, on the houses of the cooling vessels there were slots B_1 and B_2 cut. The light sources were type Kz-12-12 glimm lamps. We have performed our measurements at intensities 13,000–130,000 photons/s. The reduction of the intensity to these values did not cause a particular difficulty, since (a) we have used out a very narrow range and in this way the primary intensity was already small; (b) we blended out the light going in and coming out of the monochromator. The control of the intensity has been managed by regulating the current of the glimm lamps.

By the adjustment of the optical arrangement, special care has been taken in order that the light traveling after the mirror T be impinge completely to photocathodes of the photomultipliers. In the opposite case, a part of the light would not cause coincidences. At the same time, owing to the small light intensity and the construction of the cooling vessel, by visual observation we could not be sure about the proper adjustment. In order to get around this difficulty we have done the adjustments in the following way: we have exchanged the light source F with a large-intensity source. We have imaged the output aperture of the monochromator to the centre of the apertures B . This adjustment has been carried out using visual

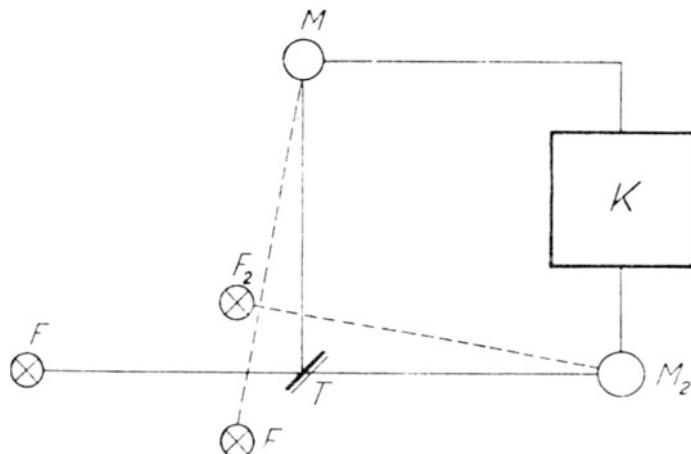


Fig. 18.4 Arrangement for measuring systematic and accidental coincidences

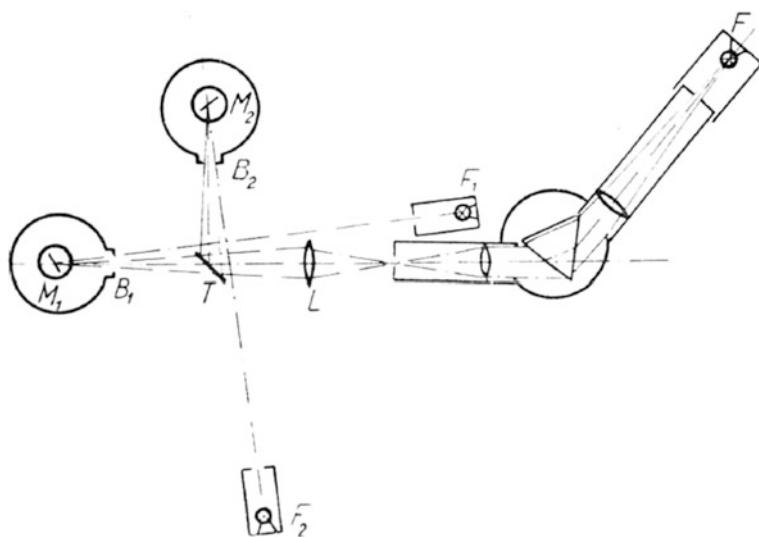


Fig. 18.5 The optical arrangement

observation. After this, we have switched to the original light source, by shifting the photomultiplier and finding the position with maximum photon count numbers. Both the photomultipliers and the cooling vessels could be moved vertically and diagonally. Besides, they could be rotated in the horizontal plane around the aperture B .

We did not need to care the positions of the light sources F_1 and F_2 , because they merely had the role of independently of each other, exciting the photomultipliers with a certain frequency.

The transmittivity and reflectivity of the mirror T have been determined with micro-lumen measuring detector. The mirror T had a transmittivity of 40% and 60% reflectivity.

18.2.5 The Counting Apparatus

The counting apparatus consisted of a photon counting device and the coincidence apparatus, described in [2, 3], respectively. On one hand, the amplified signals of the photomultipliers M_1 and M_2 have been counted, and, on the other hand, they have been coincidence-mixed (Fig. 18.6).

The coincidence signals have been counted by the end stage [“végfok”]. The coincidence mixer itself was a pentode, the signals coming from two sides has been coupled to its control and subsidiary wire meshes.

18.2.6 Control Measurements

We have thoroughly checked the electron tube part of the apparatus by using the methods described in [2, 3]. We have paid a special attention to the control of the coincidence unit, in order to see whether there are missed or erroneous coincidences. We have secured the apparatus against the external interference, however erroneous coincidences may also occur in such a way that one side of the apparatus fires due to a multiplying pulse, and the amplified signal triggers a signal in the very sensitive other side. In the control concerning such phenomena we have completely

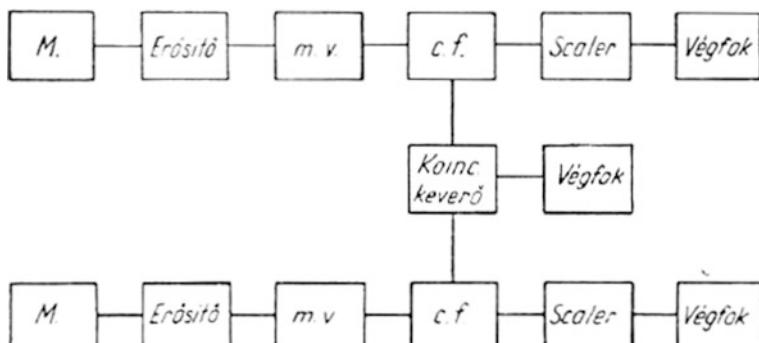


Fig. 18.6 The block diagram of the coincidence apparatus

switched on the apparatus with the electron tubes, having the electron multiplier tube also working in one side. On the other side we have interrupted the circuit of the divisor-chain of the multiplier, by using a large capacitor. In this way various disturbing pulses could reach this latter side, and erroneous coincidences could appear.

We have succeeded to completely eliminate the occurrence of erroneous coincidences.

The two sides have been adjusted so that at identical light intensity the counting will be the same. It has been an important control measurement to check whether the resolution calculated on the basis of (18.2) from the accidental coincidences and the count numbers measured in the two sides, agrees with the measured resolution capability. The latter quantity has been measured in such a way that to the coincidence mixer we have coupled in from the same photomultiplier, but one of the signals was delayed by a delay line. Beyond a suitable large delay there should no coincidences occur. The resolution time so obtained was $\tau = (2.3 + 0.2) \times 10^{-6}$ s. This value agrees well with its calculation from the accidental coincidences.

18.2.7 *The Measurements*

In the course of the measurements in each two minutes we have alternatingly measured with coherent and incoherent illuminations. In each measurement we have obtained the count numbers (N_1 and N_2) of both photomultipliers, and the coincidences. In practice, 10–50 two-minutes readings were possible by a constant count numbers. The variations in the count numbers were not large. But, still, it would have been complicated to reset to the original count numbers. It was suitable to continue the measurement in spite of the changes of the intensity, and to carry out the data processing in such a way that the comparison of the data measured by coherent and incoherent illumination, would be independent from the variation of the count numbers.

18.2.8 *The Method of Data Processing*

In order that we make the data processing independent of the changes in the intensity, it is suitable to choose the resolution time τ as basis of the data processing. This is because by the illumination with incoherent light only accidental coincidences exist [(2) formula]. Here τ is a constant characterizing the experimental apparatus, and does not depend on the intensity. The coincidence count number can be written in the form (18.2) also in the case of coherent illumination.

$$K = 2N^2 \tau^{(k)} t. \quad (18.8)$$

From (18.2) and (18.8) we have

$$\tau^{(k)} = \tau + \frac{\varepsilon \cdot p}{2N}. \quad (18.9)$$

It can be seen from (18.9) that if there are systematic coincidences, then $\tau^{(k)} > \tau$, regardless of whether the intensity is constant or not, during the measurement.

The numerical processing of the data has been carried out as follows: We have processed separately each series of measurements during which the intensity was constant, or only the allowed statistical fluctuations appeared. Then the results stemming from the individual series of measurements have been summarized. In an individual series of measurements, from the counting numbers of the two-minutes measurement of the resolution times, $\tau^{(i)}$, $\tau^{(k)}$ have been calculated for both incoherent and coherent illumination, respectively. (Henceforth the superscript indices (i) and (k) refer to data collected by incoherent and coherent illumination, respectively.)

$$\tau^{(k)} = \frac{K_\nu^{(k)} T}{2N_{1,\nu}^{(k)} N_{2,\nu}^{(k)}} \text{ and } \tau^{(i)} = \frac{K_\nu^{(i)} T}{2N_{1,\nu}^{(i)} N_{2,\nu}^{(i)}}. \quad (18.10)$$

(The subscript index ν distinguishes the measurement results obtained in the two-minutes runs, $T = 120$ s). Of course,

$$\tau^{(i)} = \tau \quad (18.11)$$

is the true resolution time.

We have calculated the averages referring to each individual series of measurement, i.e.

$$\bar{\tau}^{(k)} = \frac{\tau_1^{(k)} + \tau_2^{(k)} + \dots + \tau_m^{(k)}}{m}, \quad (18.12)$$

$$\bar{\tau}^{(i)} = \frac{\tau_1^{(i)} + \tau_2^{(i)} + \dots + \tau_m^{(i)}}{m}, \quad (18.13)$$

where m is the number of measurements within one series. In addition, we have calculated the errors of the averages:

$$\Delta \bar{\tau}^{(k)} = \left[\frac{(\tau_1^{(k)} - \bar{\tau}^{(k)})^2 + (\tau_2^{(k)} - \bar{\tau}^{(k)})^2 + \dots}{m(m-1)} \right]^{\frac{1}{2}}, \quad (18.14')$$

$$\Delta\bar{\tau} = \left[\frac{(\tau_1 - \bar{\tau})^2 + (\tau_2 - \bar{\tau})^2 + \dots}{m(m-1)} \right]^{\frac{1}{2}}. \quad (18.14'')$$

If

$$\bar{\tau}^{(k)} - \bar{\tau} < 3[(\Delta\bar{\tau}^{(k)})^2 + (\Delta\bar{\tau})^2], \quad (18.15)$$

i.e., the deviation of $\bar{\tau}^{(k)}$ from the resolution time $\bar{\tau}$ is not significant, there are systematic coincidences, or at most as many, which are compatible with the measurement error.

In order to check whether the fluctuations within one series of measurements were really merely of statistical character, we have calculated the resolution capabilities from the average counting numbers \bar{K} and \bar{N}_1, \bar{N}_2 :

$$\bar{\tau}^{(k)} = \frac{\bar{K}^{(k)} T}{2\bar{N}_1^{(k)} \bar{N}_2^{(k)}} \text{ and } \bar{\tau}^{(i)} = \frac{\bar{K}^{(i)} T}{2\bar{N}_1^{(i)} \bar{N}_2^{(i)}}, \quad (18.16)$$

where $\bar{K}^{(k)} = \frac{1}{m} \sum_1^m K_\nu^{(k)}$, etc. ... We have calculated the errors of the quantities $\bar{\tau}^{(k)}$, $\bar{\tau}$ so obtained:

$$\Delta\bar{\tau}^{(k)} = \pm \bar{\tau}^{(k)} \sqrt{\frac{1}{m\bar{K}^{(k)}}} \text{ and } \Delta\bar{\tau} = \pm \bar{\tau} \sqrt{\frac{1}{m\bar{K}}} \quad (18.17)$$

In deriving (18.17) the error of number coincidences has been expressed by the averages based on the Poisson distribution. In order to illustrate this procedure we have included the data processing of one series of measurements (Table 18.1).

18.2.9 The Results of the Measurements

We have performed altogether 119–119 two-minutes measurements, with both coherent and incoherent illuminations. In these measurements the intensity was $N = 400$ pulse/s, i.e. $n = 130,000$ photons/s.

The evaluation provided us with the data:

coherent illumination: $\bar{\tau}^{(k)} = (2.484 \pm 0.022) \mu\text{s}; \bar{\tau} = (2.484 \pm 0.022) \mu\text{s},$
incoherent illumination: $\bar{\tau} = (2.425 \pm 0.022) \mu\text{s}; \bar{\tau} = (2.434 \pm 0.023) \mu\text{s}.$

The averaging now refers to *all* the data.

It is seen that $\Delta\bar{\tau}^{(k)}$ and $\Delta\bar{\tau}$ —calculated for counter-checking—are in good agreement with the values $\Delta\bar{\tau}^{(k)}$ and $\Delta\bar{\tau}$, respectively. The corresponding errors are

Table 18.1 An example for the data evaluation for a series of measurements

$T = 120 \text{ s Kohärenz}$							(Scalar leoztas) = 256 Inkohärenz			
Sorsz	N_1l	N_2l	K	$10^9 \tau_v$	$10^{10} (\Delta \tau_v)^2$	N_1l	N_2l	K	$10^2 \tau_v$	$10^{10} (\Delta \tau_v)^2$
1	183	173	84	2429	00324	187	182	93	2502	31360
2	179	174	96	2822	11289	185	183	99	2677	01254
3	183	178	94	2642	02433	186	185	107	2847	07952
4	181	181	96	2683	03880	187	186	88	2316	06200
5	179	180	85	2415	00504	185	185	101	2702	01876
6	178	181	93	2643	02464	188	185	84	2211	12532
7	185	181	103	2816	10890	189	186	114	2969	16322
8	178	183	90	2529	00184	186	186	92	2461	01082
9	176	183	92	2615	01664	188	187	77	2005	31360
10	180	186	74	2024	21344	191	189	119	3018	20521
11	180	177	91	2615	01664	187	184	91	2421	02074
12	177	176	81	2381	01102	185	179	110	3041	22658
13	176	181	90	2587	01020	185	183	102	2758	03725
14	175	177	73	2158	10758	194	183	80	2063	25200
15	180	181	65	1827	43428	189	182	89	2369	03842
16	176	179	89	2586	01000	184	184	99	2677	01254
Σ	2866	2871	1396	39772	113948	2996	2947	1545	41037	189212
	\bar{N}_1l	\bar{N}_2l	\bar{K}	$\bar{\tau}$	$\Delta\bar{\tau} \pm 0,069 \mu\text{s}$	\bar{N}_1l	\bar{N}_2l	\bar{K}	$\bar{\tau}$	$\Delta\bar{\tau}$
	179,13	179,44	87,25	2,486 μs	$\pm 0,065 \mu\text{s}$	187,25	184,19	96,56	2,565 μs	$\pm 0,088 \mu\text{s}$
				$\bar{\bar{\tau}}$	$\Delta\bar{\tau}$			$\bar{\bar{\tau}}$	$\Delta\bar{\bar{\tau}}$	
				2,485 μs	$\pm 0,065 \mu\text{s}$			2,563 μs	$\pm 0,062 \mu\text{s}$	

also in good agreements, which shows that within one series of measurements the fluctuation was not larger than the Poisson-fluctuation. On the truth of the latter statement we can convince ourself by a direct calculation of the variances.

On the basis of (18.9), from the above data we have calculated ε and $\Delta\varepsilon$:

$$\varepsilon = \frac{2N}{p} (\bar{\tau}^{(k)} - \bar{\tau}), \quad (18.18)$$

and

$$\Delta\varepsilon = \pm \frac{2N}{p} [(\Delta\bar{\tau}^{(k)})^2 + (\Delta\bar{\tau})^2]^{\frac{1}{2}}. \quad (18.19)$$

By inserting the data from the measurements into (18.9) and (18.10), and taking into account $p = 3 \times 10^{-3}$ [2, 3], we receive the result

$$\varepsilon = 0.0076 \pm 0.0040.$$

It is seen that the value of ε is roughly of the order of its error. This means that *there are no systematic coincidences, or at most $\varepsilon + 3\Delta\varepsilon = 2\%$ of the photons may have given systematic coincidences.*

18.2.10 The Method of Frequent Changes

In the course of building the experimental apparatus it caused a lot of difficulty to secure stable count numbers. In the measurements discussed in III, we were able to manage stable count numbers in 10–50 two minutes serieses measurements. Still, we have found more suitable to work out a procedure, with the help of which we were able to make the observations completely independent from the variations of the intensity. The essence of this procedure is that we perform so frequently the changes between two measurements (with coherent and incoherent illuminations) that the fluctuations would unconditionally be negligible during one coherent and incoherent measurements. In order to surely achieve this goal we have to alternate between the two measurements more than one time in a second. This task has been solved by an apparatus with electronic tubes. The readout and collection of the count numbers stemming from the coherent and incoherent illuminations have also been solved with the help of electronic tubes. The functioning of the apparatus by the frequent changes between the coherent and incoherent illuminations is illustrated in Fig. 18.7.

The light source F , serving for the coherent illumination, is switched on by the generator A . The light sources F_1 and F_2 , serving for the incoherent illumination, is switched on by the generator B . The two generators are alternatingly working, thus, accordingly, the light source F and F_1, F_2 are alternatingly lighting. We separately

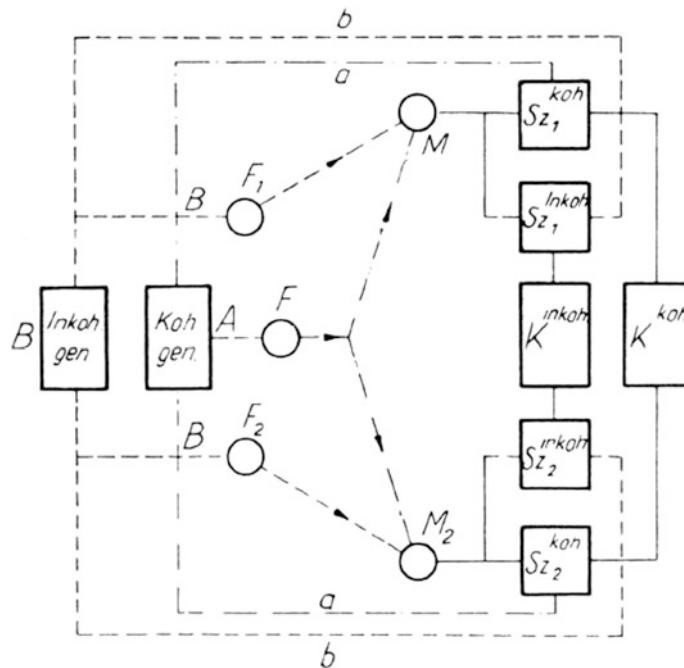


Fig. 18.7 The scheme of the apparatus by the frequent changes between the coherent and incoherent illuminations

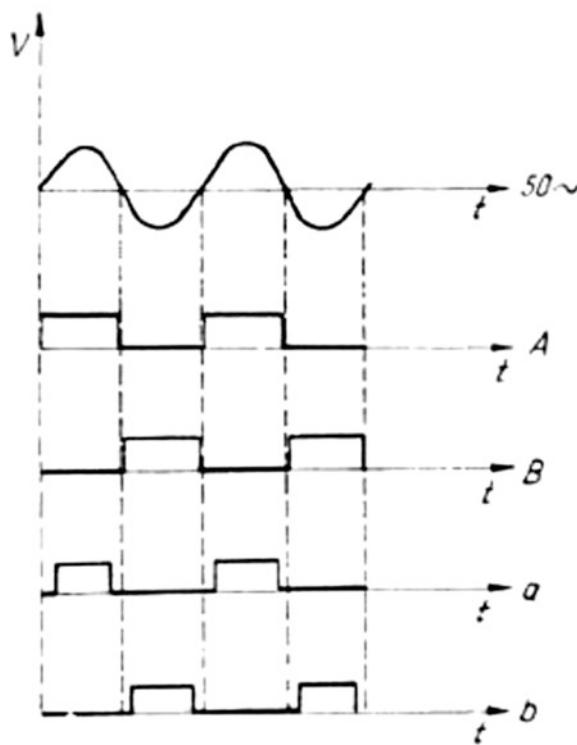
count the pulses from the photomultipliers and the coincidences stemming from the coherent and incoherent illumination, respectively. The counters Sz_1^{koh} and Sz_2^{koh} are only sensitive when the light source F is on. Similarly, the counters Sz_1^{inkoh} and Sz_2^{inkoh} are made sensitive with the signal b. The coincidence signals coming from the coherent and incoherent illumination are also counted separately.

The generator, which directed the lighting of the light sources, worked synchronously with 50 period alternating current network; giving the coherent illumination in one half-period, and the incoherent illumination in the other half-period. The temporal evolution of the signals are shown in Fig. 18.8.

18.2.11 The Controlling Measurements

The counter-checking of the apparatus has been carried out by using the method described in [2, 3]. In the processing the measurement data it is important to know, exactly how long are the commanding signals a and b. It is also essential that during the measurements the ratio of these two signals a/b should be kept constant. In order to check this latter issue, during the measurements, after each 5–5 coincidence

Fig. 18.8 The temporal evolution of controlling and the illumination signals



measurements we measured the value of a/b . This has been done by illuminating the photomultipliers with a continuously illuminating light source, and we have read out the count numbers. The ratio of the counts measured by the counters, commanded by the signals a and b , gives the value of a/b . In the evaluation we have used the averages of this ratio, and found $a/b = 0.987$. We have also measured the duration of the controlling signals, and found the value $a = 8.16 \times 10^{-3}$ s.

18.2.12 The Results of the Measurements

With the frequent changes, at $N = 300$ pulse/s $\sim 100,000$ photons/s intensities, we have carried out 306 3-min measurements. The reading out in each three minutes was needed to the statistical evaluation of the data. We have received the following results:

By coherent illumination: $\bar{\tau}^{(k)} = (2.360 \pm 0.026)$ μs ; $\bar{\bar{\tau}}^{(k)} = (2.364 \pm 0.030)$ μs ,

By incoherent illumination: $\bar{\tau} = (2.398 \pm 0.027)$ μs ; $\bar{\bar{\tau}} = (2.405 \pm 0.031)$ μs .

In these measurements, too, the values of $\Delta\bar{\tau}^{(k)}$ and $\Delta\bar{\bar{\tau}}^{(k)}$, as well as $\Delta\bar{\tau}$ and $\Delta\bar{\bar{\tau}}$, are also in good agreement.

By using the above data we have calculated ε and $\Delta\varepsilon$, and received:

$$\varepsilon = -0.0029 \pm 0.0030.$$

As can be seen, by comparing the result presented in Sect. 18.2.10, the value of ε here even turned out to be negative, thus, we truly do not need the association to it with a value differing from zero.

We have also performed measurements at smaller intensities, $N = 42$ pulse/ $s \sim 13,000$ photons/s. As a result of 159 three-minutes measurements, for ε and $\Delta\varepsilon$ we have received the value:

$$\varepsilon = -0.0017 \pm 0.0036.$$

Taking into account the results presented in Sect. 18.2.10 and in the present section, the summed-up error of ε turns out to have the value

$$\Delta\varepsilon = 0.002.$$

From this it is seen that *there are no systematic coincidences, or at most 0.6% of the photons could have given systematic coincidences*. Thus, our results are in agreement with the expectations of quantum mechanics.

* * * *

We thank Péter Faragó and Zsolt Náray, for their advices given to us during our work. We also thank our colleagues Sándor Koncz and Árpád Nagy for taking part in the planning, building and performing the measurements, as well as Ferenc Deák for the numerical calculations.

References

1. L. Jánossy, Acta Physica Hungarica **1**, 423 (1952)
2. A. Ádám, P. Varga, KFKI Rep. **2**, 313 (1954)
3. A. Ádám, P. Varga, *Magyar Fizikai Folyóirat*, under press

Chapter 19

First Observation of Photon Correlations (Bunching) with Beamsplitter and Photomultipliers (Robert Hanbury Brown and Richard Quintin Twiss)



Svetlana G. Lukishova and William J. Tango

Abstract This chapter is devoted to Robert Hanbury Brown (1916–2002) and Richard Quintin Twiss (1920–2005). Sections 19.1 and 19.2 (written by William J. Tango who worked with Twiss) are their biographies. Hanbury Brown and Twiss first carried out experiments on observation of photon correlations (bunching). Section 19.3 contains an excerpt of several pages from the book of R. Hanbury Brown “Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics” in which he describes the story of the discovery of photon correlations. The Hanbury Brown–Twiss effect marks the beginning of modern quantum optics.

In 1950s two British radioastronomers Robert Hanbury Brown and Richard Quintin Twiss observed photon correlations (bunching) using a thermal light source and an interferometer consisting of a beamsplitter and two photomultipliers. The Hanbury Brown–Twiss effect became seminal in the development of quantum optics by introducing the importance of quantum correlations [1]. In many modern quantum optics experiments for measuring photon statistics, a Hanbury Brown–Twiss interferometer remains a key part (see, e.g., Chap. 4 of the current book). The history of this discovery was described in details by one of its authors in his book [2] with some excerpts from it reproduced in Sect. 19.3 of a current book. See also a brief historical outline in [3].

S. G. Lukishova (✉)

The Institute of Optics, University of Rochester, 275 Hutchison Road, Wilmot Building, Rochester, NY 14627, USA

e-mail: sluk@lle.rochester.edu; lukishov@optics.rochester.edu; lukishova@hotmail.com

W. J. Tango

School of Physics A28, University of Sydney, Sydney, NSW 2006, Australia

e-mail: W.Tango@physics.usyd.edu.au

In the next few paragraphs the main steps of this discovery will be highlighted. It will be shown how the development of instruments for radioastronomy (for stellar angular size measurements) led to the fundamental experiment in optics which is now considered the main cornerstone in “the beginning of quantum optics as a relatively new or rejuvenated field” [1]. The Hanbury Brown–Twiss interferometer is an intensity interferometer. It does not deal with field amplitudes like a Young double slit interferometer, but rather with intensities. Conventional, two-antenna amplitude Michelson stellar interferometers, different from the Michelson interferometer for table-top optical measurements, but analogous to the Young’s double-slit interferometer, were used in the 1950s for measuring the angular sizes of radio-wave sources in the sky. They have a resolution proportional to λ/D , where λ is a wavelength and D is the separation between the antennas seen from the direction of the source. To resolve most stellar dimensions, this separation should be far beyond any available technology of the early 1950 (more than 100 km separation was needed).

The intensity signal from a radio telescope is actually noise, and Hanbury Brown had noticed in 1949 that the noise from two separate antennas appeared to be correlated. The coherence length of the intensity interferometer is orders of magnitude larger than that of the conventional amplitude interferometer. In addition, one could even record the signals received at two widely separated antennas and correlate them later. For a theory of such an intensity interferometer a mutual friend introduced Hanbury Brown to the mathematician Twiss and their joint work on the theory of the intensity interferometer appeared in 1954 [4]. Earlier, in 1952, Hanbury Brown and his students made such an interferometer [5] and measured the angular sizes of radio sources Sygnus A and Cassiopeia A. At the same time, Australian radioastronomers carried out the measurements of the angular sizes of the same sources using an amplitude interferometer which had a signal-to-noise ratio much larger than that of a Hanbury Brown intensity interferometer, so in practice an Australian amplitude radiointerferometer showed better performance.

During his 1954 visit of the Jodrell Bank observatory at the University of Manchester where Hanbury Brown worked, Twiss found that the correlations in intensity from the intensity interferometer are unaffected by ionosphere instability. Based on this remarkable property for the radiowave region, he suggested that an optical intensity interferometer also might be unaffected by turbulence effects in the atmosphere. A few months later, both Hanbury Brown and Twiss came to the very important conclusion: in spite of a large diameter of optical telescope mirrors, to increase the signal-to-noise-ratio of an intensity interferometer, this interferometer does not need precision optics, so crude “light buckets” will be sufficient. But they had concern that the correlation might not be preserved in the process of photoelectric emission. In addition, physicists with whom they discussed an optical intensity interferometer did not believe in any correlation between photons.

With these concerns, before constructing an optical intensity interferometer for measuring angular sizes of the stars, Hanbury Brown and Twiss decided on a laboratory demonstration. Hanbury Brown spent $\sim 2\text{--}3$ months assembling in a dark room a table-top version of the optical intensity interferometer. In this setup,

a very intense light source with an extremely narrow bandwidth (a mercury vapor lamp) and a half-silvered mirror as a beam splitter were used. Photomultiplier tubes served as the light detectors. The fluctuations in the output currents from the photomultipliers were amplified and multiplied together in a linear mixer. By measuring the intensity correlations between the beams in the two interferometer arms, at the different distances between path lengths in two arms, they found positive correlation at a zero relative distance and no correlation at a large distance (1.8 cm). Initially Hanbury Brown did not see any correlations, but than Twiss came and found that a narrow-band filter defocuses the image of the source reducing correlations [2].

This work was published in 1956 [6], and now this interferometer is widely used in quantum optics. Hanbury Brown and Twiss observed positive correlations at zero delay as should be for bosons, but the scientific community met their results skeptically. Even famous scientists thought that the authors were wrong and interference should not be observed between different photons. Everybody remembered Dirac suggestion [7] “Each photon interferes only with itself. Interference between two different photons never occurs”. Soon experimental results arrived from attempts to measure photon correlation using a similar interferometer, but with negative results. Some even preceded [6]. See Chap. 18 of the current book and Sect. 19.3 of this chapter with references.

The Hanbury Brown and Twiss experiment was the crucial demonstration of “photon bunching” in thermal light. Purcell in his theoretical 1956 paper [8] showed that for bosons (photons) correlations should be positive, but for electrons (fermions) they are negative. Later Glauber developed a theory of correlation functions and further developed the theory of the Hanbury Brown–Twiss effect [1]. In addition to quantum optics, Hanbury Brown–Twiss interferometers also found applications in particle physics [9]. See, for instance, [10] on measuring statistics of fermions (electrons).

After the successful table-top experiment, an optical intensity interferometer was built for measurements of the angular sizes of the stars, using correlations of photons. In 1956, Hanbury Brown and Twiss published paper [11] on measuring the diameter of the star Sirius using such an interferometer with large-diameter mirrors. They used five-foot diameter searchlights left over from the Second World War. After this success [11] they put their efforts to the larger optical intensity interferometer that was built in Narrabri, Australia with mirrors diameter of 22 feet moving on a circular railway track of diameter 618 feet with maximum resolution of 5×10^{-4} angle second [12]. Starting with this publication in Nature in 1964 [12], the angular diameters of 32 bright stars were measured in Narrabri. Using these data the temperatures of these stars were defined.

Next Sect. (19.1) of this chapter is a biography of Hanbury Brown containing excerpts from two papers: (1) Davis and Sir B. Lovell: Robert Hanbury Brown (1916–2002), *Historical Records of Australian Science* **14**, N 4, 459–483 (2003), with permission of the journal and the Australian Academy of Science; and (2) V. Radhakrishnan: Obituary. Robert Hanbury Brown, *Physics Today* **55** (7), 75–76 (2002), with permission of the American Institute of Physics. Section 19.2 is a

biography of Twiss written by William Tango who worked with him. Some information from Tango own paper [13] is included with permission of the Oxford University Press. Section 19.3 is an excerpt about discovery of the Hanbury Brown Twiss effect from the book of R. Hanbury Brown *Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics*, Radio Society of Great Britain, Bedford, UK (2016), (reprinted by courtesy of Marion Hanbury Brown and the Radio Society of Great Britain).

References

- [1] R.J. Glauber: One hundred years of light quanta, *Nobel Lecture*, December 8 (2005): https://www.nobelprize.org/nobel_prizes/physics/laureates/2005/glauber-lecture.pdf.
- [2] R. Hanbury Brown: *Boffin. A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics*, Radio Society of Great Britain, Bedford, UK (2016).
- [3] W.J. Tango: The Hanbury Brown–Twiss effect and the birth of quantum optics, *Australian Physics* **51**, N 4, 116–121 (July–August 2014).
- [4] R. Hanbury Brown and R.Q. Twiss: A new type of interferometer for use in radio astronomy, *Phil. Mag.* **45**, 663–682 (1954).
- [5] R.H. Brown, R.C. Jennison and M. K. Das Gupta: Apparent angular sizes of discrete radio sources, *Nature* **170**, N 4338, 1061–1063 (1952).
- [6] R. Hanbury Brown and R.Q. Twiss: Correlation between photons in two coherent beams of light, *Nature* **177**, 27–29 (1956).
- [7] P.A.M. Dirac: *The Principles of Quantum Mechanics*, Oxford (1930).
- [8] E.M. Purcell: The question of correlation between photons in coherent light rays, *Nature* **178**, 1449–1450 (1956).
- [9] G. Baum: The physics of Hanbury Brown–Twiss intensity interferometry: from stars to nuclear collisions, *Acta Phys. Polon. B* **29**, 1839–1884 (1998).
- [10] W.D. Oliver, J. Kim, R.C. Liu, Y. Yamamoto: Hanbury Brown and Twiss–type experiment with electrons, *Science* **284**, Issue 5412, 299–301 (1999).
- [11] R. Hanbury Brown and R.Q. Twiss: A test of a new type of stellar interferometer on Sirius, *Nature* **178**, 1046–1053 (1956).
- [12] R. Hanbury Brown, C. Hazard, J. Davis and L.R. Allen: A preliminary measurement of the angular diameter of α -Lyrae, *Nature* **201**, 1111–1112 (1964).
- [13] W. Tango: Richard Quentin Twiss 1920–2005, *Astronomy and Geophysics* **47**, 333 (2006).

19.1 Robert Hanbury Brown: Biography (1916–2002)

This section contains excerpts from two papers:

- (1) J. Davis and Sir B. Lovell: Robert Hanbury Brown (1916–2002), *Historical Records of Australian Science* **14**, N 4, 459–483 (2003), with permission of the journal and the Australian Academy of Science.

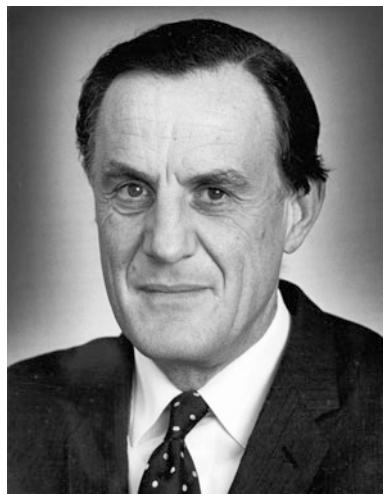
- (2) V. Radhakrishnan: Obituary. Robert Hanbury Brown, *Physics Today* **55** (7), 75–76 (2002) with permission of the American Institute of Physics.

Robert Hanbury Brown (Figs. 19.1–19.4) was born in Aruvankadu, India in 1916. He was the son of an officer in the Indian Army in India. From the age of 3 years Hanbury Brown was educated in England, initially at a School in Bexhill and then from the ages of 8 to 14 years at the Cottesmore Preparatory School in Hove, Sussex. In 1930 he entered Tonbridge School as a Judd scholar in classics [1]. His interest to science was encouraged by his grandfather, Robert Hanbury Brown, the irrigation engineer and one of the early pioneers of radio, who kept his own laboratory [2]. Hanbury Brown's interests turned to electrical engineering, and after two years he decided that he would seek more appropriate education in a technical college. Hanbury Brown decided to take an engineering course at Brighton Technical College studying for an external degree in the University of London. At the age of 19 he graduated with a first-class honors B.S., taking advanced electrical engineering, telegraphy and telephony. He then obtained a grant and in 1935 joined the postgraduate department at the City & Guilds, Imperial College, London. In 1936 he obtained the Diploma of Imperial College for a thesis on oscillators [1]. He planned to study for a Ph.D. at Imperial College, but in 1936, he drew the attention of the rector Henry Tizard, who noticed that Hanbury Brown was keener on flying airplanes than attending lectures, and recruited him to Bawdsey Manor. This fairytale castle was hidden away on the Suffolk coast, where

Fig. 19.1 R. Hanbury Brown at the Jodrell Bank Experimental Station, University of Manchester, United Kingdom (1952 year). (Reproduced by courtesy of Marion Hanbury Brown)



Fig. 19.2 R. Hanbury Brown. (Reproduced by courtesy of Marion Hanbury Brown)



Robert Watson-Watt's team was developing radar in great secrecy. Of this period Hanbury Brown later wrote, "Nothing which I have done since then has been so exciting, so absorbing, or so worth-while." [2]. In 1942, he went to Washington, DC, to liaise with the US Navy, and remained there until 1947. Hanbury Brown's career as a research scientist was remarkable [1]. When chance brought him to Jodrell Bank Observatory, University of Manchester in 1949 at the age of 33 he was known as a pioneer of radar but not as an astronomer or scientist, and he registered for the degree of Ph.D. with the aim of improving his academic qualifications. Within a few years he became a most distinguished figure in the international field of physics and astronomy. In 1960 the University of Manchester elected him to a personal chair of radio astronomy and awarded him the honorary degree of D.Sc., and he was elected a Fellow of the Royal Society of London in the same year. Hanbury Brown's immediate success in research came from the impact of his long experience in radar technology with the nascent science of radio astronomy. He was a scientist of the heroic age who could design and construct his own equipment and who seemed to thrive when faced with almost insuperable physical conditions—witness his measurements of the radio emission from Andromeda nebula in 1950 and his single-handed measurements of the angular diameter of Sirius under the appalling conditions of winter nights at Jodrell Bank in 1955 [1]. In 1962 Hanbury Brown moved to Australia to work at the University of Sydney. There he built his next interferometer—the Narrabri Stellar Intensity Interferometer—in the Australian bush, surrounded by natural and physical hazards that were antagonistic to reliable astronomical measurements, outside Narrabri in New South Wales, consisting of two 23 foot telescopes [1]. Despite only aiming to stay in Australia for two years, he stayed for a total of 27, going on to receive the highest honor of Australia, the Companion of the Order, in 1986 for his work. He made several visits to India, another country he loved. As an honorary fellow and the first Raman

Visiting Professor of the Indian Academy of Sciences, he spent time at the Raman Research Institute in Bangalore in 1974 and also toured the country [2]. He returned to the United Kingdom in the mid-1970s and in 1991 published his autobiography “Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics”.

In optics, Hanbury Brown is primarily known for his work on interferometry, in particular, in 1956, with Richard Q. Twiss, he invented the Hanbury Brown and Twiss interferometry technique—which went on to be widely used in astronomy and quantum optics [3–12]. The Hanbury Brown and Twiss effect is observation of correlations in the intensities received by two detectors from a light beam. Devices which use the effect are commonly called intensity interferometers and were originally used in astronomy, although they are now widely used in quantum optics. Hanbury Brown and Twiss designed an optical laboratory experiment on correlation measurements in 1955. A light source was formed by a small rectangular aperture, on which the image of a high-pressure mercury arc was focused. The 4358 Å line was isolated by filters, and the beam was divided by a half-silvered mirror to illuminate the cathodes of two photomultipliers. The two cathodes were at a distance of 2.65 m from the source and their areas were limited by identical rectangular apertures. In order that the degree of coherence of the two light beams might be varied, one photomultiplier was mounted on a horizontal slide that could traverse normal to the incident light. The two cathode apertures, as viewed from the source, could thus be superimposed or separated by any amount up to about three times their own width. The fluctuations in the output currents from the photomultipliers were amplified over the band 3–27 Mc/s and multiplied together in a linear mixer. The average value of the product, which was recorded on the revolution counter of an integrating motor, gave a measure of the correlation in the fluctuations. The results of this laboratory experiment, published early in 1956 [3], showed beyond question that the photons in two coherent beams of light are correlated, and that this correlation is preserved in the process of photoelectric emission. Furthermore, the quantitative results were in fair agreement with those predicted by classical electromagnetic wave theory and the correspondence principle.

Initially some opposition in the scientific community argued that Hanbury Brown and Twiss predictions violated the laws of physics. Walter Heitler and Paul Dirac repeatedly told Hanbury Brown and Twiss, that their scheme could not work [2]. See also [12] and Sect. 19.3 of this chapter. Radhakrishnan [2] also describes that at a Caltech colloquium at which Hanbury Brown talked about it, Richard Feynman jumped up and said, “It can’t work!” Hanbury Brown responded, “Yes, I know. We were told so. But we built it anyway, and it did work.” Late that night, Feynman phoned and woke Hanbury Brown up to say “you are right.” Feynman also wrote a letter in which he admitted his mistake and acknowledged the importance of this phenomenon that, at first sight, appears counterintuitive, even to quantum physics theorists.

Throughout, Hanbury Brown retained his lively sense of humour and the wider vision of a cultured man. He was internationally respected and admired—nowhere is this more evident than in the fact that within the space of a few years he

addressed the World Council of Churches in 1979 on “Faith, Science and the Future” and also presided over the International Astronomical Union meeting in Delhi during his term of office as President 1982–85. Hanbury Brown became increasingly interested in the relation of science to society. His book “The Wisdom of Science” is concerned with the relevance of science to culture and religion. His last book, “There are No Dinosaurs in the Bible”, written for his grandchildren and unpublished at the time of his death reflects his ultimate conclusion that there are fundamental issues in science that lie beyond human understanding [1].

19.2 Richard Q. Twiss: Biography (1920–2005)

Richard Q. Twiss (Figs. 19.3 and 19.4) was born in Simla in India in 1920. He was educated at Rugby School and completed the Mathematical Tripos at Cambridge with distinction in 1941. He spent the World War II years in the Admiralty working on radar, and after the war was appointed British Liaison Officer to the Research Laboratory for Electronics (RLE) at MIT in the USA, where he assisted in editing the 27-volume RLE Technical Series on radar. He was awarded a Doctor of Science degree by MIT in 1949 for his work on the theory of magnetrons. He then returned to the United Kingdom where he continued his research into the generation of electromagnetic radiation.

After the World War II, Hanbury Brown in the University of Manchester working on increasing resolution of radio interferometers, suggested a method of intensity interferometry. The analysis of the sensitivity of such an instrument was made by Richard Q. Twiss, who showed that the technique was feasible. Both



Fig. 19.3 R. Hanbury Brown leads blackboard discussion at the first Coherence Conference (Rochester, NY, USA, 27–29 June 1960). Left to right, A. T. Forrester, R. Q. Twiss, E. M. Purcell, and R. Hanbury Brown. [Reproduced with permission of University of Rochester, River Campus Libraries (University Archives)]



Fig. 19.4 R. Q. Twiss (left) and R. Hanbury Brown (right) at the first Coherence Conference (Rochester, NY, USA, 27–29 June 1960). (Reproduced by courtesy of Marion Hanbury Brown)

researchers also suggested that this method would also work at visible wavelengths. They carried out a laboratory experiment and published their results in 1956 [3–5]. Their results caused a furor. Eminent physicists said that it was nonsense and experiments purporting to refute the “Hanbury Brown–Twiss” (HBT) effect were published. One of the keystones of quantum mechanics is the Young’s double slit experiment. Interference effects can be seen when there is only one photon present at a time, but any attempt to determine which slit the photon passes through destroys the effect. In the HBT effect, photodetectors are placed behind each slit and the signals are correlated electronically. According to textbook in quantum mechanics, there should be no interference. Physicists had become so accustomed to thinking of interference in terms of single photons that they overlooked the fact that the HBT effect occurs only when large numbers of photons are present and the non-classical Bose–Einstein statistics of the photons must be taken into account. The debate surrounding the HBT effect led to a much deeper understanding of the nature of light and marks the beginning of modern quantum optics.

A Nobel laureate in physics of 2005 R. J. Glauber [13] explicitly acknowledged that he was motivated by the work of Hanbury Brown and Twiss. In 1968 Hanbury Brown and Twiss were jointly awarded the Eddington Medal of the Royal Astronomical Society for their work.

In 1955 Twiss moved to Sydney, Australia, where he took up a research position in the Commonwealth Scientific and Industrial Research Organization (CSIRO), the federal government agency for scientific research in Australia, in its Division of Radiophysics. As well as doing more work on the HBT effect, his work on

electromagnetic-wave propagation laid the theoretical foundation for both astrophysical masers (discovered much later) and the gyrotron for generating coherent microwave radiation. Hanbury Brown and Twiss initiated work on construction of the Narrabri Stellar Intensity Interferometer. Twiss selected a site (Narrabri) in New South Wales. He was appointed as senior lecturer in physics, with the intention that he would take charge of the project. Much later Richard would recall Australia as a colonial backwater and the prospect of living in rural Australia did not appeal to him. The electronics for this interferometer were made in the United Kingdom, and Twiss travelled to England to help. After a prolonged absence from Australia he resigned his position. Hanbury Brown was appointed to a chair, and with his colleagues completed and successfully operated the Narrabri Stellar Intensity Interferometer.

Twiss's calculations showed that the classical Michelson stellar amplitude interferometer had a much greater sensitivity than the intensity interferometer, but posed great technical challenges. With the rapid developments then taking place in optics he felt that it was feasible to build a modern Michelson stellar interferometer, but by now had no wish to be tied down by academic or administrative chores. He worked as a visitor in Alastair Gebbie's laboratory at the National Physical Laboratory in the United Kingdom, with funding through a grant held by Prof. Ken Smith at the University of Sussex. Prof. Walter Welford of Imperial College was largely responsible for the detailed optical design.

In 1969 the instrument was moved to the Italian field station of the Royal Observatory Edinburgh, located at Monteporzio Catone, south of Rome. Later, in 1972, it was replaced by a "Mark II" interferometer which was a major step in the development of modern optical interferometry. The Monteporzio station was closed in 1976 and Richard Q. Twiss effectively retired from active scientific research to pursue his interests in art and music. He was awarded the Albert A. Michelson Medal of the Franklin Institute in 1982, jointly with Robert Hanbury Brown.

In 1998 Richard came to Sydney for the summer opera season and visited the Sydney University Stellar Interferometer, the modern Michelson successor to the Narrabri Stellar Intensity Interferometer, also at Narrabri. Shortly before his death in Sydney on 19 May 2005, aged 84 years, he applied for Australian permanent residence. See more details of his biography in [14]. His biography will also be published in [15].

References

- [1] J. Davis and Sir B. Lovell: Robert Hanbury Brown (1916–2002), *Historical Records of Australian Science* **14**, N 4, 459–483 (2003). See also a website of Australian Academy of Sciences <https://www.science.org.au/fellowship/fellows/biographical-memoirs/robert-hanbury-brown-1916-2002>, and J. Davis and B. Lovell: Robert Hanbury Brown. 31 August 1916–16 January 2002. Elected FRS 1960, *Biographical Memoirs of Fellows of the Royal Society*, 49–83 (2003).
- [2] V. Radhakrishnan: Obituary. Robert Hanbury Brown, *Physics Today* **55** (7), 75–76 (2002).

- [3] R. Hanbury Brown and R.Q. Twiss: Correlation between photons in two coherent beams of light, *Nature* **177**, 27–29 (1956).
- [4] R. Hanbury Brown and R.Q. Twiss: A test of a new type of stellar interferometer on Sirius, *Nature*, **178**, 1046–1053 (1956).
- [5] R. Hanbury Brown and R.Q. Twiss: The question of correlation between photons in coherent light rays, *Nature* **178**, 1447–1448 (1956).
- [6] R. Hanbury Brown and R.Q. Twiss: A new type of interferometer for use in radio astronomy, *Phil. Mag.* 45, 663–682 (1954).
- [7] R. Hanbury Brown and R.Q. Twiss: The question of correlation between photons in coherent beams of light, *Nature* **179**, 1128–1129 (1957).
- [8] R.Q. Twiss, A.G. Little, and R. Hanbury Brown: Correlation between photons in coherent beams of light, detected by a coincidence counting technique, *Nature* **180**, 324–328 (1957).
- [9] R. Hanbury Brown and R.Q. Twiss: Interferometry of the intensity fluctuations in light I. Basic theory: the correlation between photons in coherent beams of radiation, *Proc. Roy. Soc. (London)*, A, **242**, 300–324 (1957).
- [10] R. Hanbury Brown and R.Q. Twiss: Interferometry of the intensity fluctuations in light II. An experimental test of the theory for partially coherent light, *Proc. Roy. Soc. (London)*, A, **243**, 291–319 (1957).
- [11] R. Hanbury Brown, *The intensity interferometer; its application to astronomy*, Wiley (1974).
- [12] R. Hanbury Brown, *Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics*, Taylor & Francis, NY (1991). Last edition: Radio Society of Great Britain, Bedford, UK (2016).
- [13] R.J. Glauber: One hundred years of light quanta, *Nobel Lecture*, December 8 (2005). https://www.nobelprize.org/nobel_prizes/physics/laureates/2005/glauber-lecture.pdf
- [14] W. Tango: Richard Quentin Twiss 1920–2005, *Astronomy and Geophysics* **47**, 333 (2006).
- [15] W. Tango: Richard Quintin Twiss, *Oxford Dictionary of National Biography*, will be published.

19.3 The Hanbury Brown–Twiss Effect

Excerpt from the book of R. Hanbury Brown *Boffin: A Personal Story of the Early Days of Radar, Radio Astronomy and Quantum Optics*, Radio Society of Great Britain, Bedford, UK (2016) (reprinted by courtesy of Marion Hanbury Brown and the Radio Society of Great Britain).

The ‘Hanbury Brown–Twiss Effect’

Our original theory had been accepted by radio engineers without a murmur because to them radio waves were simply waves and it was easy to prove by fairly

simple mathematics that our interferometer would work. However, when it came to proving that it would work with light we had to worry about photons, and there were some lingering doubts in our own minds and several well entrenched doubts in the minds of the physicists whom we consulted. What worried the physicists most was, that if the system really did work, then it was impossible to escape the conclusion that the time of arrival of photons at two separate detectors must be correlated; indeed the photons would have to arrive in pairs and they didn't like that idea at all! The correlation between photons didn't worry Richard and me at all, we simply accepted what we were told by the quantum theory; our concern was that the correlation might not be preserved in the process of photoelectric emission. In view of these doubts we decided that before going any further it would be wise to see if the thing actually did work in the laboratory.

So one day in 1955 I borrowed the dark room of the spectroheliograph at Jodrell and started to install a simple intensity interferometer. I made an artificial star by focusing a high-pressure mercury arc on a pinhole in a sheet of copper and then divided the light from this pinhole into two beams by a half-silvered mirror which illuminated two phototubes whose separation could be varied. The electrical noise in the output of the phototubes was amplified in a bandwidth of about 30 MHz and correlated in a linear multiplier.

All our research students at Jodrell were busy working on good sound radio astronomy projects and I didn't want any of them to spend their time on what might be a wild goose chase, so with a good deal of help from our workshop I put the whole thing together myself. It took two or three months but at last it was built and I spent hours and hours in the dark trying to make it work; but no matter how carefully I set it up, the correlation we expected simply wasn't there. It looked as though something was wrong with the theory-perhaps our critics were right after all.

When he heard of my troubles Richard came up from London, took a long hard look at the equipment, sang a few lines from the Erlkonig, and then pointed out a very subtle fault which I would never have noticed in a month of Sundays. After focusing an image of the mercury arc on the pinhole I had inserted a narrow band filter between the condenser lens and the pinhole and the thickness of this filter had defocused the image. As a consequence the pinhole was not acting as a true artificial star and the correlation was reduced. When this was put right the equipment worked perfectly and in a run of about 6 h I succeeded in demonstrating that the correlation between photons really does exist and that it agreed reasonably well with our calculations.

We published the results of this experiment in Nature under the title 'Correlation between Photons in Two Coherent Beams of Light' [1]. See Fig. 19.5 with schematic of experimental setup (reprinted from [1] and added to this section by the editors). In modern textbooks of Physics and Optics this correlation is called the 'Hanbury Brown-Twiss effect' and is usually regarded as the work of three people. It proved to be remarkably influential in making people think about the behaviour of light in terms of quanta; many years later when the Franklin Institute awarded us the Albert Michelson Medal in 1982 the citation read, 'it is safe to say that much of the early history of quantum optics has its roots in the Hanbury Brown-Twiss effect and that this phenomenon can rightly be viewed as a cornerstone of modern optical science'.

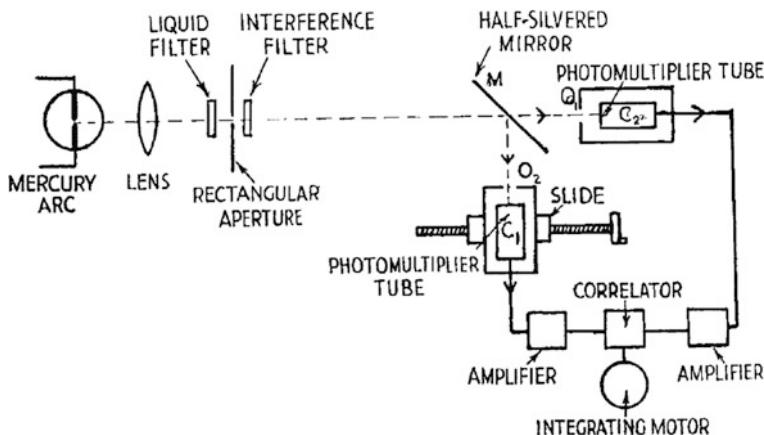


Fig. 19.5 Simplified diagram of the apparatus (added by the editors and reprinted from [1] with permission of Nature Publishing Group. See also a modern view on this experiment in R. J. Glauber: One hundred years of light quanta, Nobel Lecture, December 8 (2005): https://www.nobelprize.org/nobel_prizes/physics/laureates/2005/glauber-lecture.pdf)

Troubles with Quanta

Our work really put the cat among the pigeons. The basic trouble was that one can think about light in two different ways, as a wave or as particles. Richard and I had treated light as a wave which on arriving at the phototube causes the emission of a photoelectron. If you look at it like that then it is fairly obvious that the outputs from two phototubes will be correlated when the light falling on them is mutually coherent. However if you insist on thinking of light as a stream of independent particles like ping pong balls, which is what most physicists-especially particle physicists-prefer to do, then it is impossible to see how the arrival times of these particles can be correlated. In fact to a surprising number of people the idea that the arrival of photons at two separated detectors can ever be correlated was not only heretical but patently absurd, and they told us so in no uncertain terms, in person, by letter, in print, and by publishing the results of laboratory experiments which claimed to show that we were wrong. If science had a Pope we would have been excommunicated.

To me the most interesting thing about all this fuss was that so many physicists had failed to grasp how profoundly mysterious light really is, and were reluctant to accept the practical consequences of the fact that modern physics doesn't claim to tell us what things are like 'in themselves' but only how they 'behave'. The most common objection to our work was that the time of arrival of one photon at a detector cannot conceivably be correlated with that of another because individual photons are emitted at random times and must therefore arrive at random times. If our system was really going to work, one would have to imagine photons hanging about waiting for each other in space!

Although answers to this objection can be found in books on orthodox quantum theory, we had quite a hard job persuading people that to talk about the behaviour

of a beam of light as though it is a stream of independent photons which preserve their individual identities from emission to absorption is a gross misuse of the concept of a photon and gives the wrong answer. To think of light as a stream of photons is only useful when it is exchanging energy in quanta with a detector or absorber; but these photons cannot be pictured as preserving their individual identities in a beam of light. To describe the behaviour of light ‘on the wing’ we must use an entirely different metaphor, the wave, and these two pictures, the wave and the particle, can only be related mathematically. We had to persuade our opponents, many of whom were surprisingly irate, that there is no satisfactory mental picture of light which gives the right answer to this particular problem and that the only way of getting the right answer was to do the mathematics.

At a more sophisticated level I would be waylaid by an indignant physicist in the corridors of the University brandishing one of the sacred books of physics. Thrusting a copy of Dirac’s *Quantum Mechanics* [2] under my nose he would point to pages 14 and 15 where Dirac states that: ‘Interference between two different photons can never occur.’

He would then ask me how I thought photons could arrive in pairs if they didn’t interfere with each other. I used to point out that Dirac was writing about the photons in a beam of light which is split into two components which are then recombined so that they interfere. If he had been considering the ‘interference’ of photons in time, not space, when the two component beams are not recombined, I think he would have written something different.

There were other more complicated objections to our work some of which took an awful lot of time to answer, but at least they taught us some physics. As one example, two pundits of optics, Peter Fellgett [3] and Clark Jones, pointed out that if our theoretical treatment of the correlation between photons was correct then the classical formula for the fluctuations in the temperature of a grey body in an isothermal enclosure was wrong. Although we suspected that the classical formula itself might be wrong we couldn’t see why, and it worried us for nearly a year. Indeed we were so baffled by this problem that we wrote a letter to *Nature* defending our position by an argument which was so weak that it was immediately shot to pieces. Eventually Richard Twiss solved the problem; he realised that in classical thermodynamics the stream of radiation falling on a grey body and the stream which it emits are treated as independent which is quite incorrect because they interact. When this was put right all was well, but whether or not the formulae in the books on thermodynamics have been changed, as they should be, I have no idea.

Finally there were vigorous objections based on two laboratory experiments which claimed to show that there is no correlation between photons. The first was performed in Budapest: under the direction of an experienced physicist, Lajos Jánossy [4], who had quite recently been a member of our staff at Manchester. Jánossy illuminated two photomultipliers alternately with coherent and incoherent light and counted the coincidence of pulses produced by the arrival of individual photons in the two phototubes. In an observation lasting 10 h he found no significant increase in the coincidence rate between photons when the light on the two phototubes was mutually coherent. In a paper published in 1955 he claimed that the

experiment showed that, ‘in agreement with quantum theory the photons of two coherent light beams are independent of each other...’

The second experiment [5] was made by E. Brannen and H. I. S. Ferguson at the University of Ontario in 1956 shortly after the publication of our own laboratory experiment. Like Jánossy they looked for correlation between photons and didn’t find it; they reached the conclusion that ‘if such a correlation did exist it would call for a major revision of some of the fundamental concepts in quantum mechanics’.

The experimental procedure in both these experiments was beyond reproach but they were hopeless theoretically. The essential point, which they had both missed because they hadn’t done the necessary maths, was that the correlation between individual photons cannot be observed with a coincidence counter unless one has a very intense source of light with an extremely narrow bandwidth, such as an isotope lamp, and that is something neither of them had. They didn’t realise that Richard and I could work with a mercury arc because we were measuring the correlation with a linear multiplier which was counting about a million times more photons than their coincidence counters. We analysed both these experiments in detail and published a short note [6] pointing out that it would have taken Jánossy and his colleagues about 10¹¹ years, somewhat longer than the age of the Earth, to get a decent result! Brannen and Ferguson, so we calculated, would have had to count photons for 1000 years. All this controversy taught many physicists something new about the nature of light and was good fun from a scientific point of view, though it wasn’t good for our public relations at a time when we were trying to raise money to build a stellar intensity interferometer. In fact we decided that before we could ask anyone for the money, we must do some more experiments to establish beyond a shadow of doubt that what we were proposing to do was sound.

References

- [1] R. Hanbury Brown and R.Q. Twiss: Correlation between photons in two coherent beams of light, *Nature* **177**, 27–29 (1956).
- [2] P.A.M. Dirac: *The Principles of Quantum Mechanics*, Oxford (1930).
- [3] P.B. Fellgett: The Question of correlation between photons in coherent beams of light, *Nature* **179**, 956 (1957).
- [4] A. Ádám, L. Jánossy and P. Varga: Coincidences between photons contained in coherent light rays (in Russian), *Acta Physica Hungarica* **4**, 301–305 (1955). **Editors’ addition:** See also earlier paper of the same authors (Chap. 18.2 of the present book with translation into English) A. Ádám, L. Jánossy and P. Varga: Coincidences of photons traveling in coherent beams of light (In Hungarian; Koherens fénynyalábokban haladó fotonok koincenciái), *Magyar Fizikai Folyóirat* **2**, 499–514 (1954).
- [5] E. Brannen and H.I.S. Ferguson: Photon correlation in coherent light beams, *Nature* **178**, 481 (1956).
- [6] R. Hanbury Brown and R.Q. Twiss: The question of correlation between photons in coherent light rays, *Nature* **178**, 1447–1448 (1956).

Index

A

Active and reconfigurable metamaterials, 262
Active components, 293
Ádám, A., 590
Alignment of dot-in-rods, 152
All-optical, 289, 291, 294
All-optical signal switching/processing functions, 262
Angular position, 33
Anisotropy, 298
Antibunching, 53, 65, 67, 68, 103, 105, 106, 120, 122, 144
Antibunching histogram, 147
Anticorrelation, 3, 4
Anticorrelation experiment, 14
Anticorrelation test, 8, 12
Array, 295
Arrestin, 212
Aspect, A., 4
Asymmetric split ring metamaterial design, 262
Attenuated classical pulses, 12
Avalanche photodiode, 215
Avalanche photodiode modules, 110

B

BAPTA, 213
BB84 protocol, 104, 114
Beam splitter, 4, 17, 587, 605
Behavioral experiments, 220
Bell's inequality, 14, 53, 58, 65, 69, 71, 83
Binary logic gates, 262
Biphoton, 66, 70, 72, 85
Bistability, 293
Bloch modes, 293
Bloch sphere, 197

Boson-sampling, 187

Bowtie plasmonic nanoantennas, 132
Bragg-reflector microcavity, 128
Brannen, E., 619
Brumberg, E., 490, 511, 535
Bullseye grating, 139
Bullseye nanoantennas/gratings, 136
Bunching, 52, 60, 605
Burnham and Weinberg, 4, 13

C

Carbon nanotubes, 121
Cascaded plasmon-solitons, 303
Cascading metadevices, 253
Cat state, 53, 55, 86
Čerenkov, P.A., 493
Charge-coupled devices, 215
Circularly polarized mode of definite handedness, 128, 145–147
Circular polarized fluorescence, 145–147
Classical light sources, 224
Clauser, J.F., 4, 14
Clauser and Freedman, 4
Cluster state, 58
Coherent, 207, 275
Coherent control, 239, 243
Coherent interactions in metamaterials, 239
Coherent interactions on ultrathin films, 244
Coherent metadevices, 255
Coherent nonlinearity, 272, 305
Coherent optical gates, 262
Coherent perfect absorption in quantum regime, 261
Coherent state, 53, 64, 87
Coincidence histogram, 145

- Coincidences, 106, 590
 Collection efficiency, 123
 Color-center bulk diamond, 126–128
 Color centers in diamond, 103, 113–118,
 126–128, 149, 157
 Complex scattering matrix, 242
 Compressive direct measurement, 38, 39, 46
 Compressive sensing, 37
 Cones, 218
 Continuous variables, 57, 64, 76
 Controlling light with light, 239
 Corneal threshold, 220
 Correlations, 589
 Cyclic nucleotide cGMP, 210
- D**
 Dark adapted eye, 219
 1-D Bragg-reflector microcavities, 127
 1-D distributed Bragg-reflector microcavities,
 128
 Delayed-choice interference experiment, 21
 Delayed-choice single-photon interference
 experiments, 21
 Density matrix, 32
 De-serializer, 188
 Diamonds color centers, 103, 113–118,
 126–128, 149, 157
 Digital quantum memory, 191
 Dipolar mode, 283
 Dirac, 618
 Dirac distribution, 40, 44–46
 Direct measurement, 25, 39–41, 44, 45
 Discrete-time quantum walks, 198
 Discrete variables, 57
 2D-microcavities for epitaxial quantum dots,
 125
 Dot-in-rod, 103, 113, 152–154, 157
 2D-photonic crystal, 125
 Dynamics within the field half-cycle, 390
- E**
 Effective medium, 296
 Effective nonlinearity, 241, 242
 Effective nonlinearity in ultrathin films, 245
 Electromagnetically induced transparency, 63
 Electron density, 276
 Electron velocity, 276
 Electrophysiology, 218
 Emission diagram, 285
 Enhanced spontaneous emission, 128
 Enhancement of single-photon emission, 122
 Entanglement, 56–58, 65, 69, 72, 83, 233
 Entanglement swapping, 200, 201
 Entropy, 53, 57
- Epsilon-near-zero, 297
 Experimental anticorrelation test, 13
 Eye, 207, 481, 490, 493, 511, 555
 Eye sensitivity, 220, 481, 555
- F**
 Fabry-Perot and fiber-optics based
 microcavities, 129
 Faint optical pulses, 223
 Feeble light, 419, 422, 481, 555
 Feeble light beams, 4
 Feeble light interference experiments, 5, 419,
 422, 481
 Feedback, 75, 76, 83–85
 Feedforward, 76, 85
 Femtosecond, 300
 Ferguson, H.I.S., 619
 Fiber integrated SPSSs, 154
 Field enhancement, 269, 294
 Filamentation of ultrashort mid-infrared pulses
 in the atmosphere, 399
 First experimental observation of interference
 fringes with feeble light, 419–420,
 422–423
 First experiments on measuring light pressure,
 425–453, 455–479
 First nonlinear optical experiment, 481,
 485–489, 508–511, 543–544, 547–549,
 551–554
 First single photon sources, 3, 106–107
 Fock state, 55–57, 65, 67, 86, 189
 Four-wave mixing (FWM), 65, 66, 79, 82, 83,
 181
 Frank, I.M., 494, 498
 Free-carrier, 269, 300
 Free electron gas, 275
 Free-space communication channel, 196
 Frequency-of-seeing, 218
- G**
 Gaussian pointer, 27
 Gaussian state, 55, 72
 Ghost imaging, 70
 Giant nanocrystal quantum dots, 157
 Giant NQDs (g-NQDs), 113
 Glassy oligomeric microcavities, 144
 Glauber, R.J., 4, 6, 22–23, 52, 72, 78, 105–106,
 160, 606, 608, 613, 615, 617
 Glauber correlation function, 72, 78, 106
 Grangier and Aspect, 4
 Graph state, 58
 Greenberger-Horne-Zeilinger (GHZ) state, 58,
 62

H

Hanbury Brown, R., 605–612, 615
Hanbury Brown and Twiss intensity interferometer, 109–110
Hanbury Brown–Twiss, 106, 109–110, 589, 606–607, 613, 615–616
Hanbury Brown–Twiss effect, 8, 615–616
Harmonic generation, 280
Hecht, S., 220, 555, 558
Hecht, S., Shlaer, S., and Pirenne, M.H., 555, 564–565
Heralded and on-demand single-photon sources, 20
Heralded one-photon wave-packets, 14
Heralded single-photon pulses, 13
Heralded single-photons, 13
Heralded single photon sources, 4, 5
Heralded single-photon state, 181
Herriott cell, 184, 192
High-peak-power ultrashort pulses in the mid-infrared, 386
High-power light bullets, 386
Hilbert space, 40, 41, 56, 58, 64, 72, 197
Hong Ou and Mandel, 13
Hull, G.F., 455
Husimi function, 44, 53
Hydrodynamic model, 275, 304, 306
Hyperbolic, 298
Hyperbolic Metamaterials (HMM), 138
Hyperentanglement, 72
Hyperparametric conversion, 83

I

Icaroscope, 551, 553, 554
Indistinguishable antibunched photons, 156
Indistinguishable photons, 109
Individual rod cell, 221
Induced polarization, 271
Intensity distribution, 287
Intensity interferometers, 611
Interference fringes with feeble light, 5–6, 11, 419, 422, 492–493, 527, 529
Interference on photonic metamaterials, 239
Interferences with a single-photon, 17
Intraband, 279
Ion channels, 210
IR laser filamentation, 385
Isolated rod cells, 224

J

Jánossy, L., 587, 588, 618, 619

K

Kerr nonlinearity, 277, 288, 294, 296, 302, 305

Kerr-type nonlinearities, 272, 273
Kimble-Dagenais-Mandel, 4
Kinase, 212

L

Laser, 223
Laser filamentation experiments, 395
Laser-induced filamentation, 387, 395
Laser-matter interactions, 386
Lasing in filaments, 386
Lebedev, P.N., 425–431, 456–463, 483
Levshin, V.L., 485–486, 508
Lewis, G.N., 543–547
Lewis, G.N., Lipkin, D., and Magel, T.T., 544, 547
Light-by-light control functions, 239
Light pressure measurements, 425–453, 455–479
Linear interference, 242
Linear optics quantum computing, 187, 198
Linear polarization of single photons, 150
Liquid crystals, 149, 152
Liquid crystals and plasmonics, 153
Logic operations, 256

M

Mach-Zehnder interferometer, 17, 61
Magnetic field scattering coefficients, 252
Mandel, L., 4
Mandel's formula, 224
Memory cavity, 194
Metal–dielectric interface, 304
Metamaterials, 138, 271, 296
Metarhodopsin, 209
Metasurfaces, 243
Microcavities, 125
Mid-infrared, 385
Mid-infrared filamentation experiments, 399
Mid-infrared laser filaments in the atmosphere, 395
Mid-infrared spectral range, 385
Mid-infrared supercontinuum, 399
Mid-infrared-to-mid-ultraviolet supercontinua, 401
Mid-IR laser filamentation in the atmosphere, 386
Mixed state, 53, 57, 72, 88
Mode overlap, 282
Modulation, 291, 294
Multilayered HMM ceramics, 140
Multi octave high-energy supercontinua, 386
Multi octave sub-two-cycle supercontinua in the mid-infrared, 401
Multipartite state, 56, 57, 74, 79, 84, 85

Multi-photon sources, 187
 Multi-photon states, 189

N

Nanobeam, 127
 Nanobeam-PC cavities, 128
 Nanocavity, 125
 Nanocrystal quantum dots, 111
 Nanocrystals doped with rare-earth ions, 148
 Nano-engineered ultrathin media, 243
 Nanoparticle, 270, 282
 Nanopatterning of multilayered HMM, 139
 Nanorod, 285, 297
 Nanoscale, 268
 Nanospiral, 285
 Nanostructured, 243
 Neurons, 215
 Nichols, E.F., 455
 Nichols and Hull, 455–479
 Nitrogen-Vacancy (NV) center in diamond, 20, 21, 114–118, 126–131, 149, 157–158
 Non-classical light, 51–53, 58–61, 63–67, 70, 73, 74, 76, 83, 85–87
 Nondegenerate optical parametric oscillator, 194
 Nonlinear absorption, 289, 508
 Nonlinear coupling, 281
 Nonlinear optical device, 487, 551
 Nonlinear optical effects in phosphorescence and absorption, 547
 Nonlinear optical experiment, 508, 543, 551
 Nonlinear optics, 268, 488
 Nonlinear plasmonic crystals, 293
 Nonlinear plasmonics, 307
 Nonlinear polarization, 277
 Nonlinear polymer, 293
 Nonlinear refraction, 273, 487
 Nonlinear scattering, 287
 Nonlinear Schrödinger equation, 389
 Nonlinear SPP, 301
 Nonlinear transmission, 289
 Nonlocal, 296
 Nonlocality, 287
 Non-perturbative, 284
 NOON-state, 57, 61
 NV center nanodiamonds, 117, 157

O

O’Brien, B., 551, 553
 On-chip fabrication, 68, 83
 On-demand single-photon sources, 20

One-photon behavior, 13
 One-photon Fock state, 4
 One-photon number state, 8
 One-photon wave-packet, 9
 Optical comb, 79, 82, 83
 Optical compressor, 189
 Optical parametric chirped-pulse amplification, 396
 Optical physics in the mid-infrared, 386
 Orbital angular momentum (OAM), 29, 32–36, 41, 42, 46

P

Pairs of entangled photons, 14
 Pairs of photons, 13, 14
 Parametric amplification, 60, 62, 70, 74, 75
 Parametric down conversion, 20, 59, 61, 62, 70–75, 78, 79, 81, 87, 181
 Parametric oscillator, 57, 60, 64, 70, 75, 76, 78–80, 87
 Parametric sources of photon pairs, 20
 Patch clamp, 221
 Perturbative, 277
 Phase matching, 70, 78
 Phase space, 53, 54, 57, 58, 74, 79, 86
 Phosphodiesterase (PDE), 209
 Photocurrent in rods, 212
 Photoelectric response, 217
 Photomultipliers, 587, 605
 Photomultiplier tubes, 215
 Photon, 543, 545–546
 Photon antibunching, 128, 146
 Photon correlation, 587, 589, 605
 Photon-counting, 589
 Photon fluctuations, 222
 Photonic bandgap chiral microcavities, 142
 Photonic hypercrystal, 140
 Photon indistinguishability at room temperature, 157
 Photon-induced photocurrent, 213
 Photon noise, 222
 Photon number distribution, 222, 224
 Photon on demand, 67–69
 Photon pairs, 13
 Photons, 211, 616
 Photon statistics, 104, 207
 Photoreceptor, 207
 Photostable single emitters, 157
 Phototransduction, 214
 Pick-and-place manipulation, 155
 Pigment molecule, 209

Planar-aligned cholesteric structures, 143
Planar-alignment of liquid crystals, 149
Planar metamaterials, 243
Planar metasurfaces, 240
Planar photonic metamaterials, 239
Planar plasmonic metamaterials, 241
Plasmonic, 305
Plasmonic metamaterial, 248
Plasmonic nanoantennas, 132
Plasmonic nanogap and nanopatch antennas, 136
Plasmonic resonances, 281
Plasmonics, 268
Poissonian statistics, 222
Poisson probability distribution, 7
Polarization-to-time-bin transducer, 192, 193
Ponderomotive, 301
Ponderomotive energy, 387
Post-selection, 29
Proteins, 212
Pseudo-thermal, 207
Purcell enhancement, 126
Purcell factor, 108, 127
Pure state, 53, 55, 57, 64, 86, 88

Q

QE calibration, 232
Quadrupolar mode, 283
Quantized nature of light, 222
Quantum commitment, 64
Quantum communication, 21, 103
Quantum communication networks, 156
Quantum computer, 63, 64, 67, 79, 86
Quantum dot, 20, 63, 67–69, 87, 103
Quantum efficiency, 59, 84
Quantum efficiency (QE) of the rod cell, 230
Quantum fluctuations, 220
Quantum gates, 63, 77, 86
Quantum information, 21, 25, 36
Quantum key distribution, 55, 64, 71
Quantum light sources, 222
Quantum memory, 191, 192, 194
Quantum network, 63, 87
Quantum no-cloning theorem, 25
Quantum nondemolition measurement, 84, 85
Quantum photonic circuits, 131
Quantum pressure, 276, 287
Quantum repeater network, 200
Quantum repeater protocols based on SPSs, 156
Quantum repeaters, 156

Quantum-state tomography, 25, 26, 39, 43, 193

Quantum teleportation, 57, 71, 85

Quantum walk, 198

Quantum wave function, 26, 30, 36

Quantum zeno effect, 63

Qubit, 63–65, 87

Ququart, 64

Qutrit, 64

R

Radiative cascade, 14
Rare-earth-doped nanocrystals, 118
Rare-earth ions, 128
Rare-earth ions in nanocrystals, 157
Resonance, 283
Resonant enhancement, 288
Retina, 208, 513, 556
Retinal rod photoreceptor, 208
Rhodopsin, 208
Rhodopsin molecules, 209
Ringer medium, 213
Rod cell response, 218
Rod cells, 208, 209, 222
Room-temperature single-photon sources, 107, 157

S

Sagnac interferometer, 74, 80
Saturation of absorption, 508
Saturation of absorption and phosphorescence, 543
Saturation values for sigmoid curve, 221
Scan-free direct measurement, 41
Scattering coefficients, 246
Scattering coefficients for lossless media, 250
Scattering coefficients for nano-engineered metasurfaces, 248
Scattering matrix, 244
Schrödinger, E., 588
Second-harmonic, 304
Second harmonic generation (SHG), 284
Second-order coherence function, 105
Second order correlation function, 12, 185, 225
Second-order nonlinearity, 303
Self-focusing, 386
Self-focusing threshold, 390
Sensitivity threshold of the human eye, 209, 481, 490–492, 511, 555, 565
Separated pairs of photons, 13
Shot noise, 55, 56, 58, 60, 62, 65, 72, 81, 83, 84, 86, 87

- Shutter cavity, 194
 Signal to noise ratio, 225
 Silicon Carbide (SiC), 121
 Silicon vacancy, 114
 Single-emitter nanocrystals, 111
 Single emitters in liquid crystals, 141
 Single emitters with metamaterials, 138
 Single laser filament, 385
 Single-molecule fluorescence, 106
 Single-photon addition, 190
 Single-photon detectors, 195
 Single photon interference experiments, 3, 4
 Single-photon interference, 17, 589
 Single-photon probability, 184
 Single photon responses, 214, 230
 Single photons, 4, 214
 Single photon sensitivity of rod cells, 223
 Single-photons on-demand, 4
 Single-photon sources (SPSs), 3, 103, 107, 180, 207, 228
 Single-photon sources with definite linear polarization, 149
 Single-photon state, 6
 Single-photon subtraction, 191
 Single-photon wave-packets, 3, 6
 Singular scattering coefficients for ultrathin media, 244
 SiV, Cr and NE8 centers, 157
 SiV nanodiamonds, 117
 Slit grating, 139
 Small signal amplification, 255, 262
 Smolin state, 58
 Solitons, 302, 304
 Source of entangled photons, 4
 Sources of pairs of entangled photons, 4
 Sources of photon pairs, 5
 Sources of single-photons, 4
 Spatial multiplexing, 181
 Spontaneous emission enhancement, 108
 Spontaneous parametric down conversion, 228
 SPSs with definite linear or circular polarization, 141
 Squashing, 85
 Squeezed vacuum, 54, 57, 60, 61, 63, 72–76, 81, 86
 Squeezing, 54–58, 60, 65, 68, 73–76, 78–86, 233
 Stable single-photon emitters, 121
 Statistics of photons, 221
 Stern-Gerlach, 27
 Stokes parameters, 31
 Structural (phase) changes, 291
 Subattosecond pulse synthesis, 386
 Sub-Poissonian photon statistics, 12
 Subwavelength scale, 243
 Subwavelength thickness media, 243
 Suction electrode, 219
 Suction micropipette, 219
 Sum-frequency generation, 198
 Superconducting photodetectors, 215
 Supercontinuum spectrum, 393
 Super-Poissonian, 223
 Surface plasmon, 270
 Surface plasmon polariton, 269
 Sverdlov, Z.M., 535
 Switchable storage cavity, 183
- T**
- Tamm, I.I., 484, 494
 Taylor, G.I., 419–422
 Telecommunication, 300
 Thermal source, 222
 Thin film media, 246
 Third harmonic generation (THG), 284
 Thomson, J.J., 419–420, 422–423
 Threshold of vision for the human eye, 558
 THz bandwidth, 262
 Time-bin encoding, 196
 Time-bin qubits, 196
 Time-dependent generalized nonlinear Schrödinger equation, 404
 Time-dependent Schrödinger equation, 390
 Time-domain, 285
 Time-domain spectroscopy in the mid-infrared, 404, 407
 Time multiplexing, 180
 Timescales, 290
 Tomography, 62
 Transducin, 209
 Transduction cascade, 213
 Trivalent rare-earth ions, 118
 Twiss, Richard Quintin, 605, 611–615
 Two-dimensional hexagonal boron nitride (hBN), 121
 Two-temperature model, 278, 299
- U**
- Ultrafast, 306
 Ultrafast nonlinear optics, 385
 Ultrafast optical nonlinearities in the mid-infrared, 390
 Ultrafast photoionization, 395
 Ultrashort mid-IR pulses, 385
 Ultrathin films, 240
 Ultrathin media, 240
 Ultrathin photonic metasurfaces, 240

- Urbach, Franz, 551, 552
- V**
- Vacuum fluctuations, 59
- Varga, P., 590
- Vavilov, Sergey, 219, 481–508, 511, 531–535, 544, 555–556, 582, 585
- Vavilov—Čerenkov radiation, 493
- Vision, 560
- Visual perception, 207
- Visual system, 223
- Visual threshold, 223
- Von Neumann interaction, 26
- Vortex, 41, 42
- W**
- Wave equation, 276
- Wave-front sensing, 40, 43
- Waveguide, 292, 302
- Wave-particle duality, 3, 5, 17
- Wawilow, 482
- Weak measurements, 26, 29
- Weak value, 29–31, 33, 36, 38, 41, 42
- Weak-value amplification, 30
- Wheeler, J.A., 21
- Wheeler's delayed choice experiment, 5
- Which path information, 21
- Whispering gallery mode resonators, 64, 68, 76–79, 82, 83, 130
- Wigner function, 53–56, 68, 73, 79, 195
- W state, 58, 62
- X**
- Xenopus, 210
- Z**
- Zinc oxide (ZnO), 121
- Z-scan, 298