Pre–1940 1941–1959 1960–1974 **1975–1990** 1991–Present

OSA® | 100

1975–1990

Introduction

Michael Bass

In 1980, just 20 years after the first laser was demonstrated and about 10 years after the way to make low loss optical fibers was discovered, two miracles took place: one that lots of people noticed and that some recall and another that few noticed and that changed the course of human history. At the Winter Olympics at Lake Placid, New York, the Miracle on Ice in which the U.S.A. men's hockey team beat the much vaunted Soviet Union team was seen by tens of millions on television—lots of people noticed. However, the television broadcast of the Olympic Games, including the hockey match, was transmitted over an optical communications system using diode lasers and fiber optics. Virtually no one noticed this miracle at the time, but many billions would be affected by the technology. Optics changed the world and communications would never be the same. This section presents the pivotal events and technologies leading to optical fiber communications becoming practical.

Perhaps a few people in the mid-1970s could have foreseen that ultra-low-loss optical fibers and diode lasers would enable optics to take over the world as the dominant means of communications. Optics did just that. Not only are billions of kilometers of fiber optics communication cables in use with diode lasers as the light sources but progress continues as the demand for more and more information-carrying capacity continues to grow. New techniques for multiplexing are still being developed to enable higher throughput.

The invention of the laser and the demonstration of nonlinear optics spurred a greatly renewed interest in optics. In the period 1975–1990 that interest blossomed into many major applications and scientific breakthroughs. Nonlinear optics benefited from demonstration of excellent new materials for use in both the visible and the infrared. Periodically poled nonlinear material had been described as early as 1962 but was finally demonstrated in this period. It turned out that the periodically poled material was often a more efficient harmonic generator than its single-crystal index-matched version. These materials and greatly improved engineering made optical parametric oscillators and amplifiers available for applications requiring wavelength tunable sources. Nonlinear optics also made possible achieving ultrashort pulses, 6 picosec in this period (today 67 attosec) and supercontinuum pulses with spectral content exceeding an octave in frequency.

The list of applications of optics that developed in this period is too long to list in its entirety here. However, a few are worth mentioning because they are so common that the outstanding optics and optical design that makes them possible can be easily overlooked. They are the bar code scanner, the CD/DVD player, the laser printer, the laser pointer, the laser cut, the drilled or welded part of a finished product, the laser-marked product, the variable-focus spectacle lens, self-darkening spectacle lenses, soft contact lenses, the optical mouse, and the remote control for an appliance, as well as the display screens of televisions, computers, and mobile phones.

Between 1975 and 1990 developments of new lasers and their applications spurred demonstration of new medical innovations. The LASIK technique for vision correction based on the use of an excimer laser was developed and has now been used on ~30,000,000 patients. Optics and fiber optics have made detecting pathologies in patients more reliable and less invasive. Laparoscopic surgeries are performed today with minimal cuts because fiber optic endoscopes or miniaturized cameras can be inserted to give the surgeon vision of the problem that must be dealt with. Photodynamic therapy in which a laser is used to excite a dye that preferentially locates in tumorous tissues is another area in which optics and medical treatment have come together.

During this period spectacular progress was made in optical astronomy. The Hubble Space Telescope was launched and, after its optics were repaired, it performed spectacularly. It provided data on the content of the universe such as the number of galaxies and the presence of dark matter surrounding galaxies. Ground-based telescopes were designed and built that took advantage of adaptive optics to build large-aperture, segmented-mirror instruments that can minimize atmospheric distortions and provide superb images. These telescopes could be much larger than space telescopes and could gather more light from distant objects. Using image processing techniques and modern computers, it is now possible to link optical telescopes to greatly enlarge their effective aperture.

Whenever the field of optics is mentioned to non-optics people in the field of optics, they immediately think of their eyeglasses or contact lenses. And why not? Almost everyone will use spectacles or contacts at some point in his or her life and if they live long enough will have an implanted lens as part of cataract surgery. Progress in these areas has been remarkable. Contact lenses were invented that allow air to pass through, enabling long periods of comfortable wearing. In addition, contact lenses can now provide astigmatic correction. Spectacle lenses with continuously variable strength eliminated the need for bifocal lenses with a sharp delineation between near and distance viewing sections. Then photochromic lens materials became available enabling the wearer to no longer need different spectacles indoors and outdoors; the lenses would lighten and darken according to the ambient light environment.

By 1990 optics included light sources from continuously operating very stable lasers to lasers producing pulses as short as a few picoseconds (now a few tens of attoseconds). Optics included components small enough to be swallowed to 30-meter-diameter segmented telescope mirrors. Displays were getting so small as to be worn in a head-mounted device or so large as to be seen by 100,000 people in a stadium. Most interesting and important was that applications of optics beyond those that aid vision had become part of everyday life and so ubiquitous that most went unnoticed.

The Shift of Optics R&D Funding and Performers over the Past 100 Years

C. Martin Stickley

In the earliest days of the past century, advancements in optics were led by newly created optics companies: Kodak and its research laboratory, Bausch & Lomb, and the American Optical Company. George Eastman led the effort to found the Kodak Research Laboratory in 1912 because he saw the connection between optical science and development of new products. The Institute of Optics at the University of Rochester was not founded until 1929, after ten years of discussions. As for government, Thomas Edison urged in 1915 that a national laboratory be formed to attack issues faced by the U.S. Navy. While this resulted in the establishment of the Naval Research Laboratory in 1923, the (Physical) Optics Division was not formed until after World War II.

In July 1945 during the closing days of World War II, Vannevar Bush, the Director of the Office of Scientific Research and Development, in response to a request from President Franklin Roosevelt issued an extensive report entitled "Science—the Endless Frontier," which urged the government to establish and fund a broad program in science and applied research to fight disease, develop national security, and aid the public welfare. It urged that basic science and long-term applied research be supported in universities, that nearer-term applied research and development be funded in industry, and that military research be increased and tied to university and industry R&D programs as appropriate. It estimated the cost of this program to be $10 million at the outset rising to perhaps $50 million within five years. One of the recommendations was to create the National Science Foundation

Congress created the Office of Naval Research (ONR) in 1946 with the Naval Research Laboratory being its principal operational arm. In light of the wartime success in developing the proximity fuse, the Division of Ordnance Research was transferred from the National Bureau of Standards to create the Army's Diamond Ordnance Fuse Laboratory. The Army also created a laboratory for electronics research at Ft. Monmouth in New Jersey. The Air Force was spun out of the U.S. Army in 1947, leading to the creation of the Wright-Patterson Air Force Base Laboratories in Dayton, Ohio; the Air Force Cambridge Research Laboratory in Cambridge, Massachusetts, which had Infrared Optics as one of its major divisions; and the Air Force Weapons Laboratory in Albuquerque, New Mexico. Further, the MIT Radiation Laboratory at MIT, which was so successful during the war in radar development, was expanded and relocated near the small town of Lincoln, Massachusetts, and renamed the MIT Lincoln Laboratory. All of these played a major role in modern optics and laser development.

Corporate labs were established and grew after the war. Some of them were at GE, Bell Labs, RCA Laboratories, Hughes Research Laboratory, Westinghouse Research Laboratory, Raytheon, Texas Instruments, Perkin-Elmer, and Boeing. Figure 1 is an aerial photo of the iconic Bell Holmdel Laboratory. The growth of corporate labs was aided by fiscal help that resulted from the Vannevar Bush report and two events that accelerated the science and technology of and funding for optics dramatically: the launch of the Soviet *Sputnik* in 1957 and the demonstration of the laser in 1960.

In 1958 in direct response to *Sputnik*, President Eisenhower created the Advanced Research Projects Agency (ARPA) within the Defense Department. One of the U.S.'s limitations was a lack of broad and deep materials capability. Thus, ARPA initiated the Interdisciplinary Laboratories

▲ Fig. 1. Aerial view of Bell Holmdel Laboratory. (Courtesy of AT&T/Bell Labs.)

(IDL) program in 1960 to ensure that chemists, physicists, and electrical and mechanical engineers work together to solve the difficult research problems in materials development. This program led to the creation of the field of "materials science." The 12 universities funded in this program were MIT, Harvard, Cornell, Illinois, Stanford, University of Pennsylvania, Maryland, Brown, Chicago, Northwestern, Purdue, and University of North Carolina. A major success of the IDL program was the development of the science and technology of electronic materials, especially III-V materials such as GaAs and ternary and quaternary mixtures of them. These materials systems have been the success story of diode lasers and photonics more generally, and the scientists who went on to industrial laboratories to develop these materials systems for specific applications in optics were likely trained in one of the IDLs.

With government funding enabling universities to supply highly skilled people to industry who would lead in the revolution in optics brought on by the laser, we will concentrate on that history because it is in many ways symbolic of the transitions that took place in basic research in optics. This is not to say that other subjects such as advances in still and motion picture photography, CCD cameras, polaroid photography, electrophotographic (xerographic) copiers, laser printers, point-of-sale scanners, optical storage devices, laser machining, and optical communication systems could not show the same transitions; it is just that the laser revolution presents the changes most powerfully.

Simultaneously with the initiation of the IDL program was the demonstration of the first laser. This occurred at an industrial research laboratory using internal funds—the Hughes Research Laboratory (HRL) in Malibu, California, on 16 May 1960. As soon as other corporate labs heard in July of Ted Maiman's success, their efforts accelerated. TRG, a small company on Long Island, New York, had been funded by ARPA in 1959 to the tune of \$990,000 for laser development and is thought to be the first to duplicate Maiman's result. A number of military labs including MIT Lincoln Laboratory immediately initiated laser programs. The author was a 1st lieutenant in the U.S. Air Force at that time stationed at the Air Force Cambridge Research Laboratory (AFCRL) in Bedford, Massachusetts. He and Rudolph Bradbury had a ruby laser like Maiman's operating by November 1960. A request of \$392 was made for the purchase of capacitors and flashlamps. This request was immediately approved, as everyone was excited about the prospects of having an operating red laser!

Military labs like AFCRL, Wright-Patterson Air Force Base, and Air Force Weapons Laboratory (AFWL) typically had sufficient funding not only to fund their own projects but also to fund industrial and university proposals in areas of laser R&D that they deemed important. So the decade of the 1960s was one of intense laser activity, especially in the development of laser range finders and target designators at HRL and other companies, coherence studies of partially coherent lasers at Rochester and Brandeis and at TRG, the phenomenon of mode locking that was discovered in Nd:glass lasers by Tony DeMaria of United Technology Research Center in Connecticut, the development of parametric oscillators using $LiNbO_3$ at Bell Labs by J. Giordmaine and R. C. Miller and at Stanford by Steve Harris, the study of the dynamics of laser operation at the University of Rochester's Institute of Optics by Mike Hercher, and laser-induced damage to ruby and glass at HRL by Connie Guiliano and at American Optical Company by Charles Koester. These damage studies were funded by ARPA, but the other efforts (with the exception of the research on parametric oscillators at Bell Labs) were funded with military laboratory and ONR monies.

Meanwhile, with corporate funding at Bell Labs, Kumar Patel developed the CO_2 laser in 1964, and Joe Geusic developed the Nd:YAG laser in the same year; both lasers are still workhorses today.

▲ **Fig. 2.** Aerial view of IBM Watson Laboratory. (Courtesy of IBM Research—Zurich. Unauthorized use not permitted. Copyright owner is IBM Zurich at http://www.zurich.ibm.com/imagegallery/.)

At American Optical Company, Elias Snitzer developed the first Nd:glass rod laser as well as a Nd:glass fiber laser. About that same time, Bill Bridges of HRL achieved lasing of argon and krypton. While these achievements were extremely noteworthy, looking back at that decade, the most significant achievements for the U.S. telecommunications industry were the developments of GaAs homojunction (diode) lasers in 1962 at GE by Robert N. Hall and N. Holonyak, Jr., and at IBM by Marshall Nathan using corporate funds, and by T. M. Quist and R. J. Keyes at MIT Lincoln Laboratory, which had block funding by the U.S. Air Force. Initially, these lasers had to be cooled to liquid N_2 temperatures or below and could operate only as pulsed devices. It took the insight of Herb Kroemer of Varian Associates in Palo Alto, California, using corporate funds, to realize that if one formed a heterojunction at both sides of the homojunction where lasing was occurring, the greater bandgap at the heterojuction would prevent carrier diffusion away from the homojunction, thus leading to the first continuous-wave diode laser a year later. Kroemer received the Nobel Prize in 2000 for this achievement. Figure 2 is an aerial photo of the IBM Watson Laboratory.

With ARPA funding, Roy Paanenen at Raytheon demonstrated a 100-W argon laser that required a huge flow of cooling water. Also at Raytheon, Dave Whitehouse was the first to demonstrate a 1-kW laser with a longitudinal gas-flow CO_2 system that seemed as large as a tennis court. Ed Gerry, with ARPA funding, at AVCO/Everett Research Laboratory developed a flowing gas-dynamic CO_2 laser that had the potential for smaller size and ultra-high power because the waste heat in the gaseous medium could be removed by flowing the gas transversely out of the laser resonator. AVCO/Everett with continued ARPA funding went on to achieve very-high-power operation of the CO_2 laser as well as high-peak-power pulsed operation of rare gas lasers.

As the powers that were achieved by the CO_2 laser were high enough to fracture the “transparent” materials that were then available, a new effort had to be made to develop better optics for such lasers. Consequently, the author departed AFCRL in 1971 for ARPA to lead efforts to develop highly transparent windows and reflecting and anti-reflecting coatings. The best of the window materials that were developed were ZnSe and ZnS, and BaF_2 by Raytheon (Jim Pappis). Coating development was led by Maurice Braunstein at HRL and resulted in thorium-containing coatings with reflection coefficients exceeding 99%. Supporting university and industrial contractors were involved in these programs, with their roles ranging from modeling of optical distortions in high-power windows to development of techniques to measure absorption coefficients as low as 0.00001 cm^{-1}.

In the 1970s and 1980s, changes began to occur in the corporate world that led the corporations to reduce funding of research. First, Wall Street and the stock market expected companies to “make their numbers” on a quarterly basis as failure to do so would result in stock prices dropping. This led to corporations investing their money in the short term to the detriment of funding research that paid off mostly in the long term. Second, it was becoming apparent to management that these labs were perhaps more of a drain on profits than the corporation could afford as the research labs did not seem able to convert research results to products that would boost sales. Third, the rise of globalization meant that these companies faced competition around the world that had not mattered previously. Fourth, the U.S. Congress had initiated the Small Business Innovative Research Program to fund product development at businesses with fewer than 500 employees. Each agency of the federal government that had R&D funds was (and is) required to set aside 2.5% of these funds for such awards. In 1995 this amounted to \$950 million for product development by small businesses. While this is small compared to what U.S.

corporations spend annually for R&D, the availability of such funding attracted people to leave corporate research laboratories to develop their new ideas rather than attempt to do so in the corporate environment.

At this point, it is natural to ask, "Why weren't the research labs more efficient at developing new products?" It seems that the researchers were just not close enough to the companies' customers to know what was needed or what could be improved upon [1]. So large companies began cutting back their research laboratories in the 1970s–1990s, if not eliminating them altogether, and moving their best R&D people nearer to the front line. Instead of looking for major breakthroughs such as a laser, they concentrated instead on, as the *Economist* writes, "tinkering with today's products rather than pay researchers to think big thoughts. More often than not, firms hungry for innovation look to mergers and acquisitions with their peers, partnerships with universities, and takeovers of venture-capital-backed start-ups" [1].

The several changes mentioned above led to a shift of basic research and long-term applied research to universities and, to a smaller extent, government laboratories. While various government agencies still fund individual investigator proposals in optics, there has been a dramatic growth in Multi-University Research Initiatives (MURIs)—designed to tackle important long-range development objectives. MURIs involve universities and private companies that would be likely to commercialize the developments of the research done in the MURI. These MURIs thus take on development efforts that, 30 years ago, would have been done by a company that had its own research laboratory to perform the fundamental work necessary to develop the new product.

Reference

1. http://www.economist.com/node/8769863.

1975–1990

Through a Glass Brightly: Low-Loss Fibers for Optical Communications

Donald B. Keck

Technological breakthroughs develop through years of scientific collaboration and innovation, each discovery built upon the failures and successes of earlier work. Such was the case with the work on the first low-loss optical fiber. What began with three Corning scientists searching for a communications solution ultimately created what is now known to be a key to the Information Age.

In 1948, Claude E. Shannon [1] proved that optical carrier frequencies provided greater bandwidth than radio or microwave frequencies. But the technology of the day had not yet caught up with the science. Those looking to apply Shannon's work lacked a suitable light source, modulator, and detector technology as well as any kind of transmission conduit.

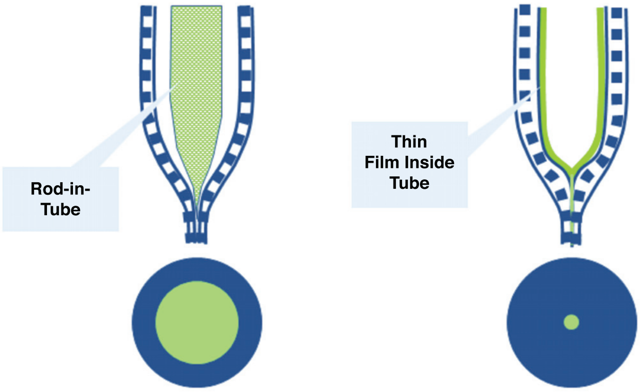
Then in 1960, Ted Maiman [2] demonstrated the first laser. A few laboratories saw it as a source for optical communications with the bandwidth that Shannon described and began to research that application. However, it could not be implemented because at that time, a suitable transmission conduit for light had not yet been invented.

Corning learned of the growing interest in optical communications on 17 June 1966, when one of its scientists, William Shaver, brought back a request from the British military. They wanted a single-mode fiber (100-μm diameter with a 0.75-μm core) with a total attenuation of less than 20 dB/km. This was prior to any publication, such as the Kao and Hockham paper [3], suggesting that optical fibers could be used as a practical communications conduit. The very best bulk optical glasses of the day had attenuation of around 1000 dB/km. The British request required an improvement in transparency of 10^{98} to reach the 20 dB/km goal. Given the science of the time, it was seemingly impossible. But within Corning's culture of scientific innovation—particularly when it came to discovering new applications for glass—"an impossible goal" was merely "a problem yet to be solved."

This particular problem was handed to Robert Maurer, a physicist known for his work on light scattering in glasses. Though Bob did not know it at the time, he actually had begun his fiber work a decade earlier. He published two definitive works in 1956 [4] and 1960 [5], indicating that Corning's flame-hydrolysis fused silica had the lowest Rayleigh scattering of all glasses he had measured.

These studies were built upon the discoveries of two giants within Corning's history, Frank Hyde [6] and Martin Nordberg [7]. In 1930, Hyde demonstrated that when vapors of silicon tetrachloride were passed through a flame in the presence of oxygen, they would hydrolyze to form a fine powder of very pure silicon dioxide that could be fused into very pure silica glass. He noted that the normal glass impurities that give rise to absorptive losses in the glass were low. Nine years later, Nordberg added titanium tetrachloride to Hyde's process and formed a very-low-expansion doped fused silica glass.

While these processes had been used at Corning for years, Bob took them in innovative directions that, ultimately, laid the foundation for the Corning group's invention of low-loss optical fiber. Always the contrarian, and influenced by his earlier work on light scattering, Bob



▲ **Fig. 1.** Illustration of RIT and thin-film processes for making an optical fiber preform. (Courtesy of Corning Incorporated.)

and a summer intern made a rod-in-tube (RIT) fiber (Fig. 1)—the best known processing method at that time—using Corning's fused silica as the cladding. He purposely added an *impurity* to the fused silica to raise the refractive index of the core, Nordberg's titanium doped silica, and obtain light guidance. Losses were still very high, but Bob was encouraged enough to request two additional scientists, Peter Schultz and Donald Keck (the author).

Peter took a fresh look at Hyde's flame hydrolysis process. He built a small boule furnace and began making various doped fused silicas and measuring their properties. Based on Bob's earlier results, the group of three focused their efforts exclusively on fused silica fibers made by flame hydrolysis. They continued the counterintuitive approach, adding an impurity to the pure fused silica to raise the refractive index and create the fiber core.

So began a time of trial and error. No human endeavor progresses more rapidly than can be measured. The group began to systematically measure and identify the sources of their optical losses. They knew absorptive losses were one source, and they struggled to examine the impurities introduced in the flame hydrolysis glasses that could cause absorption. The best analytic equipment of the day could measure impurity levels only to the parts-per-million level, and parts-per-billion were needed. An attempt was also made to evaluate losses in a few centimeters of bulk glass, but this still could not produce the losses in an actual fiber that had gone through all the processing steps. Making their own fibers was the only way to get a thorough understanding of optical losses.

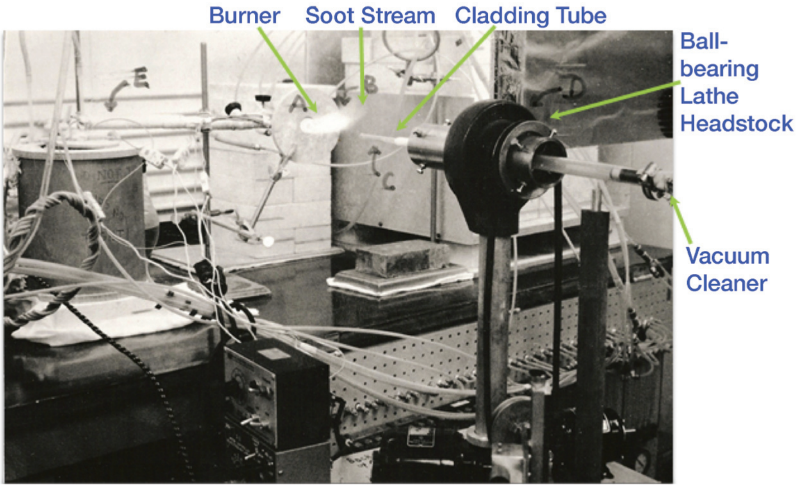
Optical absorption from formation of reduced-titanium (Ti^{3+}) color centers during the high-temperature fiber drawing step accounted for about half of the fiber loss. At first the losses were annealed away by heat-treating the fibers at 800°C to 1200°C. Unfortunately this treatment drastically weakened the fibers as a result of surface crystallization. The other half of the loss originated from light-scattering defects at the core–cladding interface. No publication of the day ever mentioned this most significant source of loss. The Corning group believed that this loss originated during the RIT process from dirt in the lab environment.

With each failure a little more was learned until an idea was hit upon that proved to be the key: the traditional RIT method was abandoned and a new approach was invented. Rather than inserting a core rod, the group decided to directly deposit a thin layer of core glass inside a carefully flame-polished cladding tube (Fig. 1). This produced intimate contact between core and clad materials and, it was hoped, would get rid of the scattering defects observed in the RIT fiber.

For those who believe that excellent work can be done only with the very latest equipment, take note of the Corning lab pictured in Fig. 2. The equipment was crude but effective. A portable lathe headstock held the rotating cladding tube in front of the flame hydrolysis burner. The burner produced a soot stream containing titania-doped silica. Initially the soot would not go into the 5–6-mm hole in our cladding tube. One of the group spotted the lab vacuum cleaner. Putting this at the end of the cladding tube beautifully sucked soot from the flame and deposited a uniformly thin layer onto the inside tube surface. This coated tube was then placed in the fiber draw furnace where the soot sintered into a clear glass layer, the hole collapsed to form a solid rod containing the doped core, and the entire structure was drawn down into fiber.

Measuring that first low-loss fiber was an unforgettable experience. It was late afternoon, and, after heat-treating a piece of the group's latest fiber, the author positioned it in the attenuation measurement apparatus. With a viewing telescope he could observe and position the focused He-Ne laser beam on the fiber end. When the laser beam hit the fiber core, a blindingly bright returning laser beam was produced. It took a moment to realize that the laser was being retro-reflected off the far end of the fiber and coming back through the optical system.

Fig. 2. Photograph of apparatus for making the first low-loss optical fiber. (Courtesy of Corning Incorporated.)



The brilliant laser beam emanating from the end of the fiber was so dramatically different from anything previously seen that it was apparent something special had occurred. With considerable anticipation, the author measured the fiber loss, and to his delight and surprise it was ~17 dB/km. With little sense of history, Donald Keck's excitement was registered in his now fairly well-known lab-book entry: "Whoopee!" (Fig. 3).

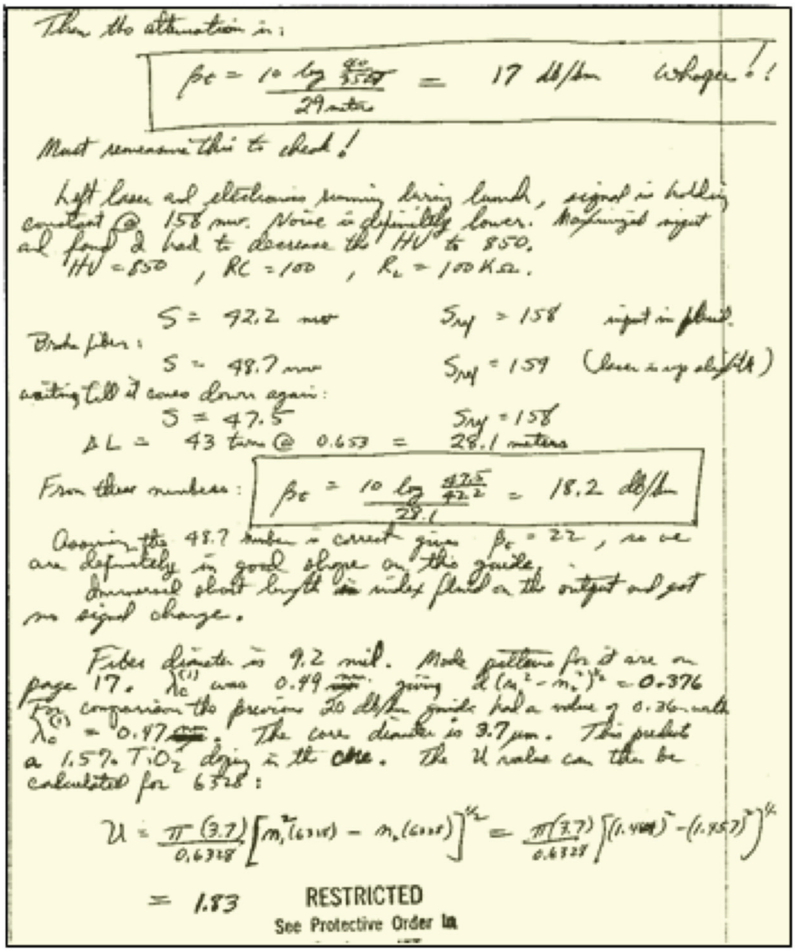
In 1970 the result was announced to the world when Bob presented the Corning group's paper "Bending losses in single-mode fibers" at an Institution of Electrical Engineers Conference in London on analog microwave technology [8]. In that paper, he mentioned that the fiber had a total attenuation of only 17 dB/km, prompting scientists at the conference to remark that at least their 2-in. helical microwave guides could be filled with lots of optical fibers. We also submitted our paper to *Applied Physics Letters*, and it was initially rejected! The reviewer commented, "It is rather difficult to visualize an amorphous solid with scattering losses below 20 decibels per kilometer, much less the total attenuation." Eventually, however, the paper was published [9]. (See Fig. 4.)

The Corning group had done it, but they were far from done. Though revolutionary, their breakthrough fiber solution was not exactly robust. Only small preforms could be made, and the heat treatment required to achieve low attenuation made the fibers brittle. Also, the preferred fiber design had shifted to multi- rather than single-mode. The larger core diameter was believed necessary to more easily couple light into the fiber from the relatively crude semiconductor lasers of the day.

To make such fibers, Peter, our colleague Frank Zimar, and the author invented another flame hydrolysis approach later dubbed "outside vapor deposition." In this method, first core and then cladding soot were deposited onto a removable rotating rod to build up a porous soot preform. Because of the lower temperature in this process, Peter found he could incorporate new dopants that had vaporized in the higher-temperature boule process. One of these dopants was germania, a glass former like silica.

In June 1972, the first fiber incorporating germania was drawn in the core. The group was obviously on the right track, as the bright light of the draw furnace was still visible through the end of a kilometer of fiber on the wind-up drum. The loss measured was only 4 dB/km, no heat treatment was needed, and fiber strength was excellent. This was the first truly practical low low-loss fiber.

This writing marks the 42nd anniversary of the Corning group's invention of low-loss optical fiber. With more than 1.6 *billion* kilometers of it wrapped around the globe, a world has been created that is dependent upon reliable, speed-of-light access to people and information anywhere, anytime, through almost any device of their choosing. The dramatic increase in users has brought with it unprecedented demand for bandwidth. Several sources, including a University of Minnesota Internet Traffic Study and Cisco, have estimated that the average Internet traffic today worldwide is ~150 Tb/s and growing at about 50% per year.



Then the attenuation is:

$$\boxed{\beta_t = \frac{10 \log \frac{40}{35.8}}{29 \text{ meters}} = 17 \text{ dB/km} \quad \text{Whoopee!!}}$$

Must remeasure this to check!

Left laser and electronics running during lunch, signal is holding constant @ 158 mv. Noise is definitely lower. Maximized signal and found I had to decrease the HV to 850.
HV = 850, RC = 100, R_L = 100 KΩ.

S = 42.2 mv $\quad$ S_{ref} = 158 signal in fluid.
Broke fiber:
S = 48.7 mv $\quad$ S_{ref} = 159 (laser is up slightly)
waiting till it comes down again:
S = 47.5 $\quad$ S_{ref} = 158
ΔL = 43 turns @ 0.653 = 28.1 meters

From these numbers: $\boxed{\beta_t = \frac{10 \log \frac{47.5}{42.2}}{28.1} = 18.2 \text{ dB/km}}$

Assuming the 48.7 number is correct gives β_t = 22, so we are definitely in good shape on this guide.
Immersed short length in index fluid on the output and got no signal change.

Fiber diameter is 9.2 mil. Mode patterns for it are on page 17. $\lambda_c^{(1)}$ was 0.49 μm giving $a(n_1^2 - n_2^2)^{1/2} = 0.376$. For comparison the previous 20 dB/km guide had a value of 0.36 with $\lambda_c^{(1)}$ = 0.47 μm. The core diameter is 3.7 μm. This predicts a 1.5% TiO_2 doping in the core. The U value can then be calculated for 6328:

$$U = \frac{\pi(3.7)}{0.6328}\left[n_1^2(6328) - n_2^2(6328)\right]^{1/2} = \frac{\pi(3.7)}{0.6328}\left[(1.4\text{[illegible]})^2 - (1.457)^2\right]^{1/2}$$

$$= 1.83$$

RESTRICTED
See Protective Order In

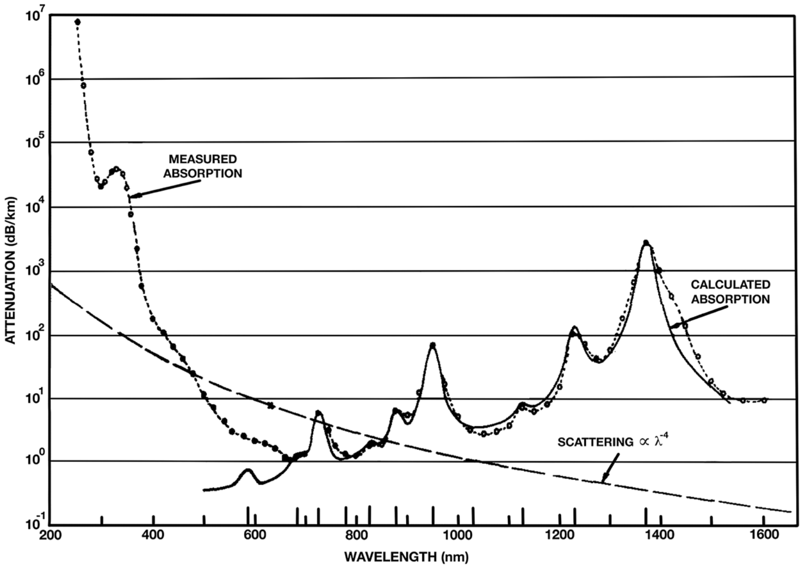
Fig. 3. Laboratory notebook with the first sub-20-dB/km fiber measurement. (Courtesy of Corning Incorporated.)

This growth rate is not surprising. Collectively we have moved from simple audio to increasing video content in our communications. Estimates are that two-thirds of the mobile data traffic will be video by 2015 as social networking continues to explode. People sending data is one thing, but machines-talking-to-machines (M2M) as is happening increasingly is yet another. The latter will overtake the former in just two or three years—all this without even considering potential new data-generating applications. We are already seeing the deployment of fiber-enabled remote sensors to monitor our environment. Power lines and highway and civil structure monitors provide an optical fiber safety net supporting the infrastructure we rely upon every day. Emerging biomedicine and biotechnology applications ranging from transmission of x-ray data to real-time high-definition video for remote surgeries to the potential petabytes involved in DNA data transmission and analysis are still in the future. It is now well established that creative people will invent new ways to use the "bits" if technology can provide improved "cost of transmitting the bit."

The amount of information that can be transmitted over a single fiber today is staggering. Commercial core networks today operate at 50 Tb/s on a single fiber, and as reported at OFC 2012, scientists are achieving in their labs record data rates of more than 305 Tb/s.

While this capacity is enormous, fiber bandwidth is finite—perhaps only 10 times higher than today's core network traffic level. Our current demand for bandwidth will most likely exceed our capacity before 2030. This would require a beginning over-build of the core networks even as we finish the build-out of the local loop! We should not be surprised if the 1.6-billion-kilometer fiber network of today will be but a fraction of that which will exist in just a couple of decades.

▶ **Fig. 4.** *Applied Physics Letters* paper [10] on the ultimate fiber losses and predicting that a 0.2 dB/km loss would be possible near 1550 nm. (Reproduced with permission from D. Keck, R. Maurer, and P. Schultz, Appl. Phys. Lett. **22**, 307 (1973). © 1973, AIP Publishing LLC.)



But beyond all the bits and bytes, the most important story of the communications revolution brought about by optical fiber may well be the one about improving human lives. All of us who have worked and continue to work in optical fiber communications technology have truly made the world a better place—and for that we should be proud.

When asked about glass, most people *still* picture something breakable that shatters when dropped. But low-loss optical fiber has shown us that hair-thin strands of glass filled with light are strong enough to help people all over the world shatter long-held assumptions and break down centuries-old political and cultural walls.

In 2000, the United Nations created the Millennium Project, aimed at lifting millions of people in the developing world from impoverishment, illness, and death. One of the primary methods for achieving that objective was to deploy the benefits of optical fiber technology for their education and economic betterment.

The International Telecommunications Union continues to track progress toward that end. In 2011 they reported that today, thanks to optical fiber, more than two billion people around the world are instantaneously and simultaneously accessing the Internet, virtually 75% of the world's rural population has cell phone coverage, and more than 60% of the world's countries have a National Research and Education network.

We have come a long way since we first stood on the shoulders of those giants of early optical communications. Today the optical fiber network has become the lifeblood of our society, providing the medium through which commerce and culture are being simultaneously created and communicated on a personal and global scale. We can never be sure just what the future of optical communications holds, but given the remarkable history of low-loss fiber, it is fairly certain to be a future full of light.

References

1. C. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, 1949).
2. T. H. Maiman, "Stimulated optical radiation in ruby," Nature **187**, 493–404 (1960).
3. C. Kao and G. Hockham, "Dielectric-fibre surface waveguides for optical frequencies," Proc. IEE **113**, 1151–1158 (1966).
4. R. D. Maurer, "Light scattering by neutron irradiated silica," J. Phys. Chem. Solids **17**, 44–51 (1960).
5. R. D. Maurer, "Light scattering by glasses," J. Chem. Phys. **25**, 1206–1209 (1956).

6. J. F. Hyde, "Method of making a transparent article of silica," U.S. patent 2,272,342 (10 February 1942).
7. M. Nordberg, "Glass having an expansion lower than that of silica," U.S. patent 2,326,059 (3 August 1943).
8. F. P. Kapron, D. Keck, and R. D. Maurer, "Bending losses in single-mode fibers," in *IEE Conference on Trunk Telecommunication by Guided Waves*, London, 1970.
9. F. P. Kapron, D. Keck, and R. Maurer, "Radiation losses in glass optical waveguides," Appl. Phys. Lett. **17**, 423–425 (1970).
10. D. Keck, R. Maurer, and P. Schultz, "On the ultimate lower limit of attenuation in glass optical waveguides," Appl. Phys. Lett. **22**, 307–309 (1973).

Erbium-Doped Fiber Amplifier: From Flashlamps and Crystal Fibers to 10-Tb/s Communication

Michel Digonnet

The deployment of the world's optical telecommunication network starting in the 1980s was a major change of paradigm in modern society that enabled the Information Age. From a technical standpoint, of the many technologies without which this colossal achievement would have never seen the light of day—from frequency-stable laser sources to efficient low-noise detectors, division wavelength multiplexers, optical filters, and low-noise high-speed electronics—perhaps none was as decisive and challenging as the fiber-optic amplifier (FOA) in general, and the erbium-doped fiber amplifier (EDFA) in particular. Like the optical fiber itself, the EDFA had no good alternative; had it not existed, no other component would have been available, then or now, to perform its vital function as nearly perfectly as it does.

The basic idea of transmitting data encoded on light carried by optical fibers dates back to at least the 1960s. Early incarnations of optical communication links used electronic repeaters that periodically detected, amplified, and remodulated the traveling light signals. Such repeaters worked adequately for high-speed communications over planetary distances, but they required power and costly high-speed electronics. By then the potential of replacing them with optical amplifiers, devices that would amplify the modulated signals without the need for electronics, had already been formulated. Optical amplifiers already existed, and they offered, at least on paper, multiple advantages, including an unprecedented bandwidth in the multiterahertz range. Yet it took nearly three decades of gradually intensifying research in numerous laboratories around the world to turn this concept into a reality, which involved, among other things, developing a practical optical amplifier utilizing a fiber as the gain medium.

From the start, the development of FOAs was riddled with challenges. To be successful in a communication network, an amplifier had to meet tough criteria. It had to provide a high, nearly wavelength independent gain over a broad spectral range while also incorporating an efficient means of mixing the excitation source with the incoming signal, being internally energy efficient, preserving the single-mode character of the trunk fiber, and inducing negligible crosstalk between channels. In later years other requirements were added to this list that further complicated the task. In retrospect, it is easy to trivialize the now well-known solutions to these problems. But back in the 1970s and 1980s when these problems were being tackled, there was nothing obvious about them, and, as in other scientific pursuits, many potential solutions were proposed, tested, and discarded.

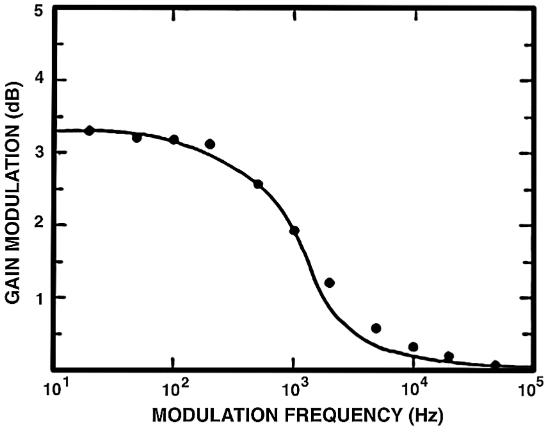
The first report of amplification in a fiber appeared in a famous article published in 1964 by Charles Koester and Elias Snitzer in The Optical Society's (OSA's) *Applied Optics*, just four years after the demonstration of the first laser [1]. This historic amplifier consisted in a 1-m Nd-doped glass fiber coiled around a pulsed flashlamp and end-probed with 1.06-μm pulses. This visionary device already contained several of the key elements of modern FOAs, including a clad glass fiber doped with a trivalent rare earth, an optical pump, and means of reducing reflections from the fiber ends to avoid lasing. It provided a small-signal gain as large as 47 dB, which is remarkable considering that it came out so early in the history of modern photonics. For his many contributions to the fields of FOAs and lasers, Elias Snitzer was awarded the OSA's Charles H. Townes Award in 1991 and the John Tyndall Award in 1994.

Like almost all the laser devices of the time, this fiber amplifier was side-pumped: the pump was incident on the fiber transversally. This made the device bulky, inefficient, and ultimately impractical. The concept of a fiber amplifier in which the pump is end-coupled into the fiber emerged years later as part of efforts carried out at Stanford University to develop a compact fiber amplifier. This work involved end-pumping Nd-doped crystal fibers with an argon-ion laser. This work demonstrated that end-pumping could produce sizeable gain (~5 dB) from a very short fiber (~cm).

The second key improvement was the introduction of the wavelength-division-multiplexing (WDM) coupler to mix the pump and the signal and end-couple them simultaneously into the gain fiber. The advantages of this technique were overwhelming: it made it possible to efficiently inject, with a compact and mechanically stable device, both the pump and the signal into the gain medium. It took several years before it was adopted, in part because commercial WDM couplers were almost nonexistent. It is now the standard technique used in the vast majority of FOAs.

Another concept critical to the performance of FOAs in general, and bench-tested first with EDFAs, is that they should use a single-mode fiber. Although in recent years new findings have suggested that the data transmission capacity could be increased by using multimode fibers, in current telecommunication links a single-mode FOA offers two key advantages, namely, a higher gain per unit pump power due to the higher pump intensity and the elimination of modal coupling, which would otherwise induce time-dependent losses at the trunk fiber/FOA interfaces.

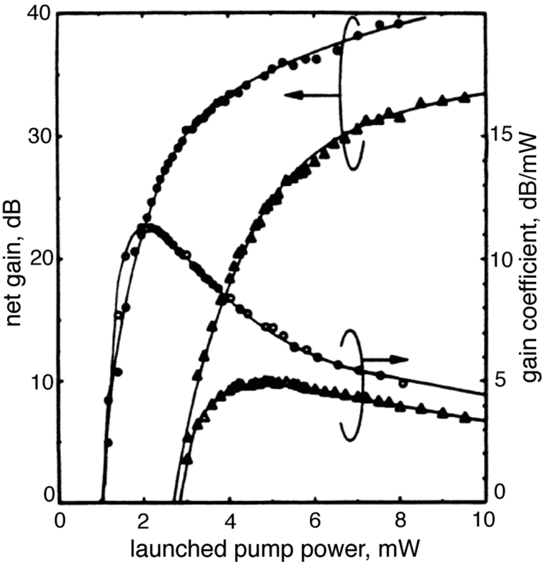
It was known as early as the 1980s that the third communication window, centered around 1550 nm, was the most promising candidate for long-haul fiber links, because in this spectral range both the loss and dispersion of conventional single-mode silica fibers are minimum. Trivalent erbium ions (Er^{3+}) in a variety of amorphous and crystalline hosts had also long been known to provide gain in this wavelength range, so this ion was a natural candidate. David Payne, who would receive the OSA's John Tyndall Award in 1994 for his pioneering work on EDFAs, and his team at the University of Southampton were first to demonstrate this potential experimentally with the report of the first EDFA in 1987 [2,3]. This was followed later the same year by a similar paper from Bell Laboratories. These milestone publications provided experimental proof that single-pass gains exceeding 20 dB were readily attainable in single-mode Er-doped fibers (EDFs) end-pumped with the best laser wavelengths available at the time, namely, 670 nm and 514.5 nm. Another key property that made Er^{3+} so attractive is that the lifetime of its 1550-nm transition is unusually long (~5–10 ms); hence the population inversion and the gain essentially do not respond dynamically at the very high modulation frequencies of the signals (unlike semiconductor amplifiers). The important consequence is that the crosstalk between signals being amplified simultaneously in an EDFA can be exceedingly small (see Fig. 1) [4], a crucial property for communications.



▲ Fig. 1. Measured crosstalk between two channels, characterized by the peak-to-peak gain variation induced in a first signal (channel A) by a second signal (channel B) sinusoidally modulated at frequency *f*. [C. R. Giles, E. Desurvire, and J. R. Simpson, Opt. Lett. **14**, 880–882 (1989)].

The EDFA seemed to be a great candidate, but several issues, some perceived to be critical by the communication community, made it difficult to be accepted right away. In fact, it took nearly another decade of detailed engineering and the development of several parallel technologies (diode lasers and fused WDM couplers, in particular) to make this device a reality.

To be practical, the EDFA had to be pumped with a semiconductor laser. Over time several pump sources and wavelengths were investigated. This battle was one of the most technically challenging and interesting in the history of the EDFA. The proliferation of inexpensive GaAs diode lasers in the 800-nm range in the electronic products of the late 1980s led to substantial research on 800-nm pumping, in particular at British Telecom Research Laboratories. However, they found that the gain efficiency was low. The reason was later identified as the unfortunate presence of excited-state absorption around 800 nm in Er^{3+}, a limitation that could not be sufficiently reduced by adjusting the pump wavelength or the glass composition.



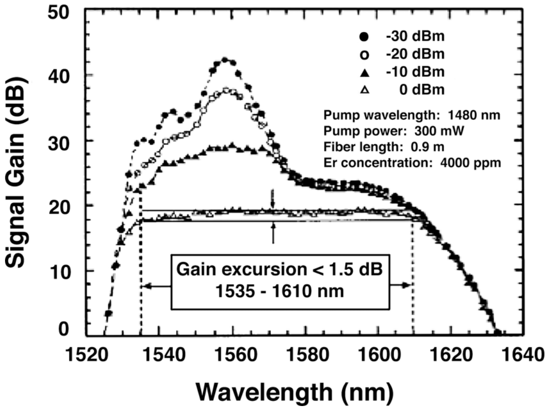
▲ Fig. 2. Measured gain and gain coefficient in an EDFA pumped at 980 nm and 1480 nm. Circles, 980 nm; triangles, 1480 nm. (Reproduced with permission of the Institution of Engineering and Technology.)

Much of this research soon focused on two pump wavelengths only, namely, 980 nm from the then-emerging strained GaAlAs laser technology and around 1480 nm from InGaAsP diode lasers. The prevalent thinking was initially that since an EDFA pumped at 1480 nm is nearly a two-level laser system, it should be difficult to invert and exhibit a poor noise performance. The demonstration of the first EDFA pumped at 1.49 μm by Elias Snitzer in 1988 quickly changed this perception. The following year saw the first report of an EDFA pumped at 1480 nm with an InGaAsP diode laser, at NTT Optical Communication Laboratory in Japan. This spectacular result (12.5 dB of gain for 16 mW of absorbed pump power) put the EDFA on a new track by establishing that a packaged FOA was within reach. For a short while pumping at 980 nm was the underdog, in part because it had a higher quantum defect than 1480-nm pumping, hence an expected lower efficiency, and in part because of the lower maturity of the strained GaAlAs technology. But 980-nm pumping nevertheless eventually won. Stimulated emission at 1480 nm turned out to be a serious penalty, which gave a lower pump efficiency and noise performance than with a 980-nm pump. M. Shimizu and his team at NTT illustrated this compromise clearly in a cornerstone paper [5] that compared the gain of an EDFA pumped at either wavelength (Fig. 2). The gain and the gain per unit pump power (11 dB/mW!) were all substantially higher with 980-nm pumping, and the transparency threshold was lower. This new understanding triggered a substantial R&D effort in the semiconductor laser community, which ultimately lead to the commercialization of reliable, high-power, long-lifetime diode lasers at 980 nm.

Many other important engineering issues were addressed through the mid-1990s. Two teams contributed to this major effort more prominently than any other, namely, David Payne's group at the University of Southampton [2,3] and Emmanuel Desurvive, first at the AT&T Bell Laboratories, then at Columbia University and Alcatel in France [6]. Many other substantial contributions came out of academic and industrial laboratories around the world, especially in the U.S., UK, Denmark, Japan, and France.

A significant fraction of the research was consumed by the quest for ever greater gain bandwidth, lower noise, and a gain that is nearly independent of signal polarization, signal power, and number of channels. To increase the bandwidth a number of ingenious solutions were implemented, ranging from hybrid EDFs concatenating fibers of different compositions and slightly offset gain spectra to adjusting the level of inversion to produce preferential gain in the C band (1530–1565 nm) or L band (1565–1625 nm) or designing the EDF so that it does not guide well above 1530 nm to produce efficient gain in the S band (1460–1530 nm). This effort was greatly complicated by the parallel need for a uniform gain (or flat gain spectrum) so that all channels have a similar power and signal-to-noise ratio at the receiver. Here too, clever solutions were conceived, from using passive filters to hybrid EDFAs, gain clamping, and the use of telluride fibers. This last approach produced a gain with a remarkable bandwidth of 80 nm (see Fig. 3) [7]. Later refinements produced EDFAs with a gain flatness well under 1 nm over wide bandwidths [8].

The EDFA rose from the status of research device to stardom remarkably rapidly, a resounding manifestation of its practical importance, exceptional performance, and timeliness. The first commercial EDFA appeared in 1992. By 1998 over 40 companies were selling EDFAs; the count ultimately



▲ **Fig. 3.** Measured gain spectrum of a 0.9-m-long tellurite EDFA at various input-signal power levels. The gain in the 1535–1570-nm range was compressed by using higher-power input signals. [Y. Ohishi, A. Mori, M. Yamada, H. Ono, Y. Nishida, and K. Oikawa, Opt. Lett. **23**, 274–276 (1998)].

peaked above 100. Research on communication systems followed suit, leading to the demonstration of increasingly large and high-performance experimental and deployed systems. As one of many examples illustrating the phenomenal performance of communication links utilizing EDFAs, in a particular experiment a total of 365 signals were simultaneously recirculated 13 times around a ~500-km fiber loop containing ten EDFAs (one every ~50 km). At the output the power imbalance between channels was as low as -7 dB and the bit-error rate only 10^{-13}. This system accomplished a remarkable total optical reach of 6850 km and a total capacity as high as 3.65 Tb/s. Deployed links now exceed 10 Tb/s over even longer distances.

In the early 2000s, following the saturation of the telecommunication industry and the sharp decline in the world's markets, a significant percentage of the optical communication task force redirected its vast technical expertise to other areas of photonics. This concerted effort gave the EDFA and other FOAs a second carrier in spectacular new applications, especially fiber sensors and high-power fiber lasers. Using an FOA to amplify the output of a fiber laser, in a now widely used configuration called the master-oscillator power amplifier (MOPA), turned out to be the most energy-efficient way to produce extremely clean and spectrally pure laser outputs up to enormous power levels. Today, fiber MOPAs are the world's brightest light sources, to a large extent thanks to the superb properties of fiber amplifiers. Yb^{3+}, in particular, rapidly became the workhorse of high-power fiber lasers for its low quantum defect and high quenching-free concentration. Power scaling posed significant challenges, including efficient coupling into the gain fiber of the high required pump powers, wavelength conversion due to stimulated Brillouin and Raman scattering, optical damage, and photodarkening. These challenges were met with a number of clever engineering solutions, including large-mode-area fibers (in which the signal intensity, and hence nonlinear effects and optical damage, are reduced) and acoustic anti-guiding fibers (in which the spatial overlap between acoustic and optical modes, and hence the nonlinearity, are reduced). Commercial fiber lasers utilizing MOPA configurations now offer average powers up to the 100-kW range, a feat that would not have been possible without the superb attributes of FOAs.

References

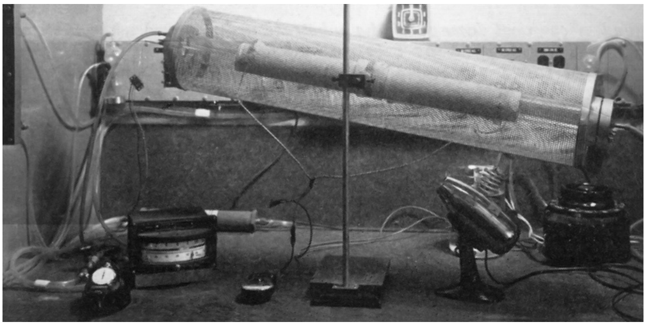
1. C. J. Koester and E. Snitzer, "Amplification in a fiber laser," Appl. Opt. **3**, 1182–1186 (1964).
2. R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "High gain rare-earth doped fibre amplifier operating at 1.55 μm," Proceedings of OFC, Reno, Nevada, 1987.
3. R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," Electron. Lett. **23**, 1026–1028 (1987).
4. C. R. Giles, E. Desurvire, and J. R. Simpson, "Transient gain and cross talk in erbium-doped fiber amplifiers," Opt. Lett. **14**, 880–882 (1989).
5. M. Shimizu, M. Yamada, M. Horiguchi, T. Takeshita, and M. Okayasu, "Erbium-doped fibre amplifiers with an extremely high gain coefficient of 11.0 dB/mW," Electron. Lett. **26**, 1641–1643 (1990).
6. M. Desurvire, *Erbium-Doped Fiber Amplifiers—Principles and Applications* (Wiley, 1994).
7. Y. Ohishi, A. Mori, M. Yamada, H. Ono, Y. Nishida, and K. Oikawa, "Gain characteristics of tellurite-based erbium-doped fiber amplifiers for 1.5-mm broadband amplification," Opt. Lett. **23**, 274–276 (1998).
8. M. J. F. Digonnet, *Rare-Earth-Doped Fiber Lasers and Amplifiers* (Marcel Dekker, 2001).

Advent of Continuous-Wave Room-Temperature Operation of Diode Lasers

Michael Ettenberg

The story of getting to room temperature continuous-wave (CW) operation of semiconductor diode lasers will start when the author arrived at RCA Labs with a fresh Ph.D. in June of 1969. RCA had decided that GaAs would be the next important semiconducting material in the solid-state electronics business after germanium, which at the time was the most prevalent transistor material. While silicon transistors were already being manufactured, GaAs transistors would be far superior, and RCA Research Lab researchers would concentrate their efforts on GaAs and related compounds to leapfrog silicon. The choice had some validity. GaAs was a direct-band semiconductor and thus had shorter electron hole lifetimes and a larger bandgap, making possible transistors with higher speed, less temperature dependence, higher operational temperature ranges, and smaller size. While all this is true, silicon became the pervasive electronic device material for a variety of good reasons that will not be detailed here. But GaAs and its related direct bandgap materials could do something that silicon could not do, that is, emit light efficiently. So RCA Labs moved its GaAs efforts to develop LEDs and diode lasers.

The author's first assignment was to grow AlAs epitaxially on GaAs single-crystal substrates via vapor-phase epitaxy, where Al is transported by passing HCl gas over Al and As is supplied by breaking down arsine. After AlAs growth characterization, devices became of interest. Since it seemed easier to make devices out of new materials than to create the materials themselves, the author joined a group headed by Henry Kressel working on laser diodes. These small devices were fascinating, as they were able to put out large amounts of reasonably directed light, albeit they could be seen only with a night vision scope. There were four relatively large research efforts at the time: Bell Labs the largest by far, Standard Telecommunications Laboratory (STL) in England, RCA Labs, and the Russian effort, about which less was known, mainly due to the cold war. At IBM, GE, and Lincoln Labs, even though diode lasers were first demonstrated there, research efforts were not substantial. The research projects at Bell Labs and STL were considerable, supported by telephone usage; telephone companies were utilities at time. The telephone giants first saw lasers as a potential source for free-space communications, a secondary effort compared with microwave transmission in air and pipes until Charles Kao envisioned optical communications in fibers [1] and research at Corning demonstrated low-loss optical fibers in 1970 [2]. Then the laser efforts intensified. RCA's research was driven by other applications such as optical disc recording and playback and military usage. Since RCA had an Aerospace and Defense division, the diode laser efforts could be justified, but RCA as a corporation was focused on television, and lasers were not a mainline effort. The research was about half supported by the corporation and half by government research contracts. The first applications of laser diodes were military in nature, and RCA decided to make the devices commercially so they could supply them to their defense customers and potentially lower the price by supplying them for other commercial uses. In 1969 RCA became the first commercial supplier of laser diodes, although it was a miniscule business, especially for a multi-billion-dollar corporation.



▲ **Fig. 1.** First liquid-phase epitaxy (LPE) growth apparatus for creating laser diode (tipping furnace). (H. Nelson, RCA Rev. **24**, 603 [1963]. Courtesy of Alexander Magoun.)

The author was introduced to diode lasers by Herb Nelson, who invented liquid-phase epitaxy (LPE) the process used to fabricate lasers throughout their initial development, well past the first CW demonstrations and many years beyond, through the first several years of CD player manufacture. Today almost all lasers are made by metal organic chemical vapor deposition (MOCVD): a much better controlled process and one that can be readily scaled up to multiple large wafers. What Herb demonstrated was called a tipping furnace (as shown in Fig. 1); it was a tubular furnace about six inches in diameter mounted in a metal cage. The cage was in turn mounted in a seesaw arrangement at the center of the furnace so the furnace could be rocked back and forth or tipped. Inside the furnace was a sealed quartz tube with hydrogen flowing through it and a carbon boat; at one end of the boat was a small polished single-crystal GaAs wafer of about a square centimeter, and at the other end of the boat was a polycrystalline GaAs wafer with a glob of gallium on it. The process started with the furnace being heated to about 800°C with the polycrystalline GaAs and Ga side lower, and some time was allowed so the GaAs could go into solution in the Ga until saturation; then the furnace was tipped the other way and the saturated Ga rolled onto the single-crystal wafer. Next the furnace was cooled and the GaAs in solution precipitated onto wafer to form an epitaxial layer on the single crystal. This epitaxial layer was much superior in terms of contaminants and defects to the underlying substrate and was also superior in terms of its luminescent properties and ability to make lower-threshold, more efficient lasers. Al was added to the Ga glob so that AlGaAs alloys could be grown. Al and Ga atoms are about the same size; therefore dislocations caused by lattice parameter mismatch would not be formed when AlGaAs was grown on GaAs. In addition, adding Al to GaAs raised the bandgap and lowered the index of refraction of the alloy compared to GaAs, which proved to be crucially important to the creation of low-threshold efficient lasers. The temperature was controlled by hand, using a variable transformer to control the current to the furnace and a 0–1000°C dial thermocouple readout. It was amazing how such a crude growth system could produce such sophisticated devices, but Herb understood the materials and was an artist. Later the process was brought under better control using carbon boats with a wafer that slid under multiple bins to allow growth of multiple layers with controlled composition and remarkable submicrometer-thickness accuracy.

The first diode lasers were simply millimeter-sized cubes of GaAs containing a diffused p–n junction with polished faces for mirrors that operated at liquid nitrogen temperatures with multi-amp very-low-duty-cycle short pulses applied. It was remarkable that these devices lased, considering that all prior laser types required tens of centimeters of cavity length and mirror reflectivity greater than 95%. The gain in GaAs per unit length was exceptional, thus allowing the gain to exceed the loss even though the reflectivity at the natural mirror surface of GaAs is only about 30%. The applied current to these first devices was many tens of thousands of amperes per square centimeter at liquid nitrogen temperatures; the threshold current increased exponentially as the temperature increased, so room temperature CW lasing was a long way away. It was found early on that GaAs cleaved nicely on the 100-crystal plane, so the lasers were grown on single-crystal wafers cut on the 100 plane. Then the mirror facets could be easily formed after the wafers had been thinned, metalized on the n and p sides by cleaving into bars. Next the bars were sawed into individual dies about 400 μm in length between the mirrors and 100 μm wide. The sawn roughened sides prevented lasing from occurring crosswise to the mirrors.

There were three important steps to room-temperature CW lasing: the addition of heterojunctions to the laser structure, the double heterojunction, and finally, the stripe contact. In 1969 independently and simultaneously, Kressel and Nelson [3] and Panish *et al.* [4] published papers demonstrating that adding heterojunctions of AlGaAs on GaAs and diffusing the p–n junction a micrometer or so from the

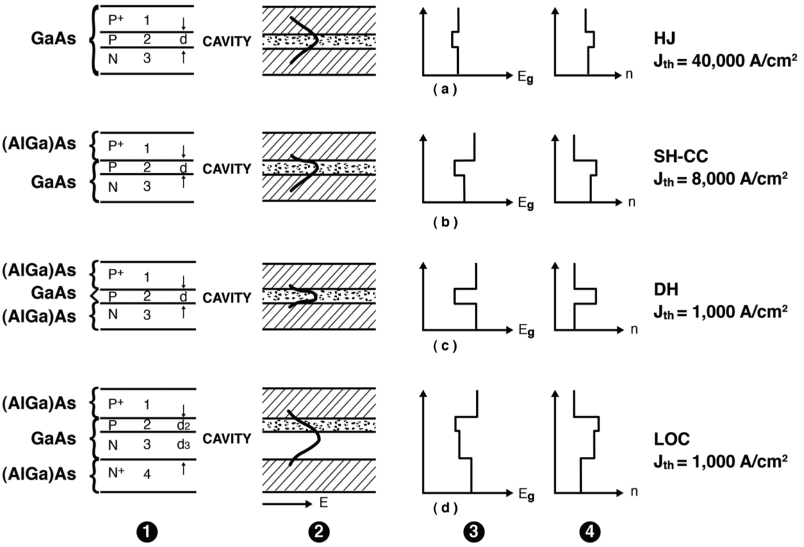


Fig. 2. Schematic cross section of various laser structures showing the electric field distribution E in the active region, variation of the bandgap energy Eg, and variation of the refractive index *n* at the lasing photon energy.
(a) Homojunction laser made by liquid phase epitaxy, (b) single-heterojunction "close-confined laser," (c) double-heterojunction laser, and (d) large-optical-cavity (LOC) laser. Figure 5 from H. Kressel, H. F. Lockwood, I. Ladany, M. Ettenberg, Opt. Eng. **13**, 417–422, 1974. (©1974 SPIE reprint_permission@spie.org.)

heterojunction formed a light waveguide. This waveguide confined the light, creating electron/hole recombination to that waveguide as illustrated in Fig. 2 [5]; consequently, the threshold current could be reduced to about 10,000 amps/cm^2, still a factor of 10 or so away from what would be needed for CW operation. The reduction in threshold from the simple *p-n* homojunctions came from the fact that the light and the recombination of electron and holes was confined to a smaller volume, thus requiring less current to invert the population to the point of lasing. These single-heterojunction devices were the first laser diodes to go into production, becoming optical proximity sensors for the sidewinder missile.

Art D'Asaro and colleagues [6, 7] at Bell Labs developed the stripe contact, shown in Fig. 3, which is a necessary and enduring feature for laser diodes, because it not only stops the cross lasing in a simple manner but facilitates the heat sinking of the device with unpumped regions all along the laser cavity.

The final and most important step came from Alferov *et al.* [8]. Alferov was one of the leaders in the field and came to United States to visit RCA and Bell Labs, among others. The visit was memorable, because it was very strange. We sat in a small office and discussed the progress of lasers. Alferov had a large heavyset man with him who said very little and seemed to know little about lasers; it was surmised that he was KGB. Alferov did not disclose the double-heterojunction laser structure nor was he shown much because the work was partially supported by Department of Defense. It was learned later that he did discuss the double-heterojunction work at Bell Labs, probably because they were more open. There was a race to achieve CW operation. Bell and RCA Labs were neck and neck, but Bell had the stripe-contact technology and learned about the double-hetrojunction structure. As a result, Hayashi and

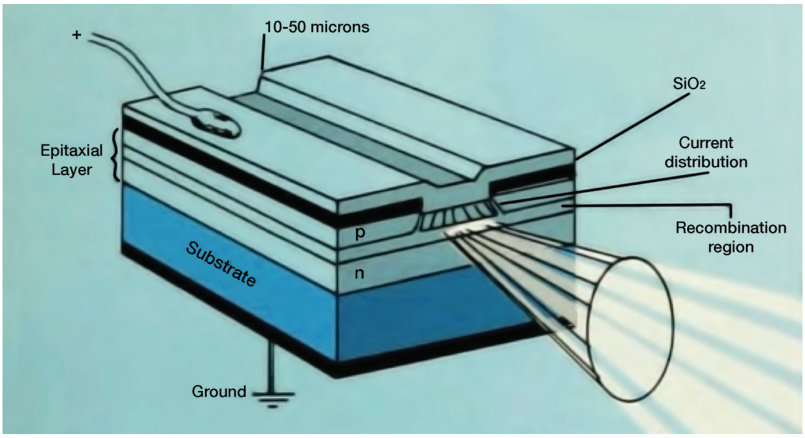


Fig. 3. Schematic of a typical CW heterojunction laser, drawn upside down to show the stripe contact. Diffraction causes the vertical spreading of the beam. Reproduced with permission from Fig. 6 of H. Kressel, I. Ladany, M. Ettenberg, and H. Lockwood, Physics Today **29**(5), 38 (1976). (Copyright 1976, American Institute of Physics.)

Panish [9] won the race to CW. The addition of the second heterojunction forced the light and electron/hole recombination to be confined to a few tenths of a micrometer and allowed thresholds close to 1000 A/cm^2, which together with the stripe-heat sinking allowed CW operation at room temperature. But it was CW in name only. The initial devices lasted only minutes.

To be useful the devices had to live for many thousands of hours, and here the author was able to make a contribution. One of his first projects on lasers came from a suggestion by Herb Nelson. He said they were evaporating SiO_2 followed by gold as mirrors on the back facet of the devices to make them emit out of one end; many of the devices were shorting probably due to pinholes in the oxide. Could the process be improved? A multi-layer dichroic reflector was eventually developed consisting of Si and Al_2O_3 as the reflector and an Al_2O_3 passivating and reflectivity control layer on the emitting facet [10]. The lifetime of AlGaAs lasers operating at low power was steadily increased to more than a million hours median time to failure [11, 12] by growing on low-dislocation substrates to eliminate defects inside the laser and applying the aforementioned passivating optical coatings to the emitting facets. Such devices helped create the early fiber-optic communications systems and were the light sources for CD and DVD players.

The final steps to today's modern laser diode were separately confining the light and the electron/hole recombination first described by Lockwood *et al.* [13] as shown in Fig. 2 and the understanding by Yariv *et al.* [14] that by making the electron/hole recombination very thin (a few tens of nanometers), quantum effects would come into play and the gain would substantially exceed what might be expected for such thin layers. The changes allowed the threshold current to be reduced to close to 100 A/cm^2, allowing lasers to be fabricated with electricity-to-light conversion efficiencies exceeding 75%. These lasers, called separate confinement heterojunction quantum well lasers, are the most reliable and efficient light sources known to man and continue to change our world.

References

1. C. Kao, IEEE Meeting in London, 1966.
2. F. P. Kapron, D. B. Keck, and R. D. Maurer, "Radiation losses in glass optical waveguides," Appl. Phys. Lett. **17**, 423–425 (1970).
3. H. Kressel and H. Nelson, "Close-confinement gallium arsenide P-N junction lasers with reduced optical loss at room temperature," RCA Rev. **30**, 106–113 (1969).
4. M. B. Panish, I. Hayashi, and S. Sumski, "A technique for the preparation of low-threshold room-temperature GaAs laser diode structures," IEEE J. Quantum Electron. **5**, 210 (1969).
5. Figure 5 from H. Kressel, H. F. Lockwood, I. Ladany, and M. Ettenberg, "Heterojunction laser diodes for room temperature operation," Opt. Eng. **13**(5), 417–422 (1974).
6. Figure 6 from H. Kreseel, I. Ladany, M. Ettenberg, and H. Lockwood, "Light sources," Physics Today **29**(5), 38–42 (1976).
7. J. E. Ripper, J. C. Dyment, L. A. D'Asaro, and T. L. Paoli, "Stripe-geometry double heterostructure junction lasers: mode structure and cw operation above room temperature," Appl. Phys. Lett. **18**, 155–157 (1971).
8. Zh. I. Alferov, V. M. Andreev, E. L. Portnoi, and M. K. Trukan, "AlAs-GaAs heterojunction injection lasers with a low room-temperature threshold," Sov. Phys. Semicond. **3**, 1107–1110 (1970).
9. I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumski, "Junction lasers which operate continuously at room temperature," Appl. Phys. Lett. **17**, 109–110 (1970).
10. M. Ettenberg, "A new dielectric facet reflector for semiconductor lasers," Appl. Phys. Lett. **32**, 724–725 (1978).
11. R. L. Hartman, N. E. Schumaker, and R. W. Dixon, "Continuously operated (Al, Ga)As double-heterostructure lasers with 70 °C lifetimes as long as two years," J. Appl. Phys. **31**, 756 (1977).
12. M. Ettenberg, Electron. Lett. **14**, 615 (1978).
13. H. F. Lockwood, H. Kressel, H. S. Sommers, Jr., and F. Z. Hawrylo, "An efficient large optical cavity injection laser," Appl. Phys. Lett. **17**, 499–501 (1970).
14. P. L. Derry, A. Yariv, K. Y. Lau, N. Bar-Chaim, K. Lee, and J. Rosenberg, "Ultralow-threshold graded-index separate-confinement single quantum well buried heterostructure (Al, Ga)As lasers with high reflectivity coatings," Appl. Phys. Lett. **50**, 1773–1775 (1987).

Remembering the Million Hour Laser

Richard W. Dixon

In the late 1960s, Bell Labs had a problem. The nation's demand for long-distance telecommunications services was steadily increasing, but the technologies then in use—coaxial cable and point-to-point microwave transmission through the air—could not keep up with the pace. The major reductions in optical fiber waveguide losses reported in the early 1970s were therefore of great interest. The lowest-loss regions of these fibers were in the 0.8 to 0.9 μm range, which could in principle be accessed by devices built using the GaAs-GaAlAs material system. Thought was given to the possible use of GaAs light-emitting diodes (LEDs), but it was immediately obvious that semiconductor lasers would be much better sources—if they could be developed reliably in commercial quantities. One could easily imagine an efficient GaAs laser that could couple a milliwatt of optical power into a fiber with a core diameter of about 50 μm. Thus was defined the first generation of fiber-optic telecommunication systems.

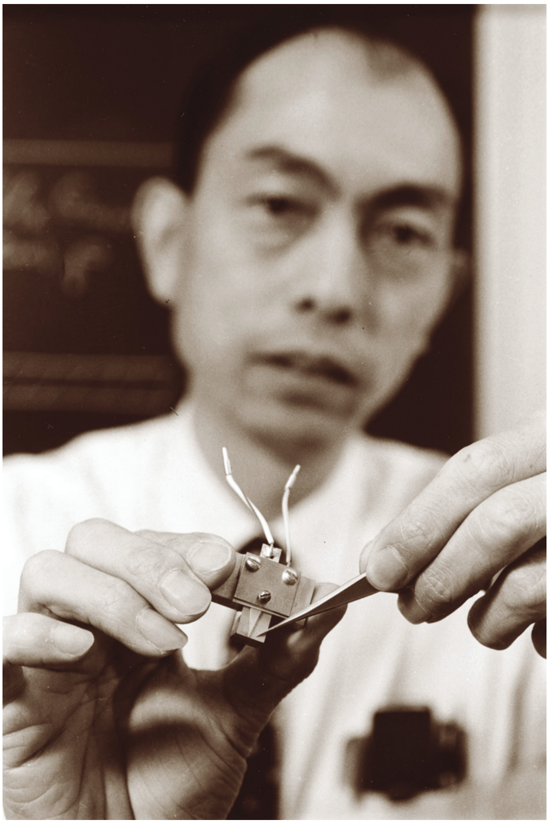
In the late 1960s and early 1970s, the author was a young supervisor working on the development of LEDs for Bell System applications. In that process, he learned quite a bit about the physics, technology, and transfer-to-volume manufacture of III-V semiconductors. One result of this program was the successful implementation of green-emitting GaP LEDs for nighttime dial illumination in the handset of the Dreyfuss-designed Trimline phone. Something like 100 million of these sets were subsequently produced.

In 1973, the author transferred to a small exploratory development group working on semiconductor lasers. The group had benefited from an excellent research effort that happened just down the hall. Most notable was the demonstration in 1970 of a continuously operating room-temperature GaAs-AlGaAs heterostructure semiconductor laser [1] (see Fig. 1). However, these broad-area lasers had high operating currents (around 400 mA) and very short lives (they were sometimes referred to as flashbulbs), but they showed the way forward!

The group's choice of a laser structure for initial development consisted of four planar epitaxial layers grown sequentially by liquid-phase epitaxy (LPE) on a GaAs substrate. We inhibited lateral carrier flow by using proton bombardment to define a "stripe-geometry" wherein only a narrow stripe, 10 × 250 μm, was electrically pumped (see Fig. 2). These "stripe-geometry lasers" became the workhorses of the early Bell Labs semiconductor laser development. They allowed the sorting out of many reliability and device performance issues. In a typical week, half a dozen or so wafers were processed into some thousands of lasers. Fast turnaround made it possible to quickly and systematically iterate device, processing, and material innovations.

Many of the early stripe-geometry lasers had very erratic properties. Some would lase for a time but would then suddenly become inoperable. Others would die slowly. Still others would not work from the outset. Typical continuous-wave operating lifetimes at room temperature were on the order of minutes to days. Many devices also had other undesirable characteristics, for example, nonlinear light output versus current. It was clear that the group had a very difficult development project on its hands! Some thoughtful observers, including one key Bell Laboratories vice president, opined that success was unattainble.

Important clues to improvements came in early 1973 from an experiment in which "windows" were fabricated on the substrate side of stripe-geometry lasers in such a way that spontaneous emission (and scattered stimulated emission if present) from the stripe region of the laser could be observed with an infrared optical microscope. Dark-line defects (DLDs), which



▲ Fig. 1. Izuo Hayashi, holding a heat absorbing device, points to the location of a broad-area semiconductor laser designed by Bell Laboratories scientists. (Bell Laboratories/Alcatel-Lucent USA Inc., courtesy AIP Emilio Segre Visual Archives, Hecht Collection.)

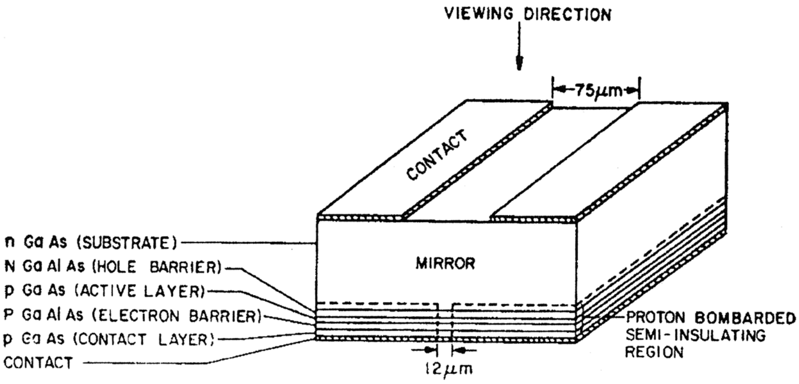
grew in a laser's active region during operation, were observed and were determined to be the principal failure mechanism in devices that stopped working in the first 100 hours or so [2].

This paper correctly stated that "the combination of low-strain processes and extreme cleanliness in materials growth should provide a dramatic increase in laser life." It galvanized a large technical community such that it seemed that everyone in the world with an electron microscope then decided to investigate this area. A picture was, in this case, worth many thousand words!

The Bell group subsequently worked hard to understand and eliminate localized modes of degradation, including those associated with DLDs in the long narrow-lasing region of the laser and those associated with mirror surfaces. Subsequent experiments showed that DLDs identical to those seen in lasers could be generated by optical pumping of undoped and unprocessed laser material, thus confirming that DLD initiation and growth could result from properties of laser material that were not associated with proton bombardment, *p-n* junction dopants, or contact metallization technology.

Many improvements in LPE growth technology and its automation were also made during this period. Fundamental difficulties with this "batch" process made it stubbornly difficult to reproducibly control, but it was greatly improved in the skilled hands of the Bell group's crystal growers.

By late 1974, with continuing work on many technology fronts, the reliability situation had improved considerably, and selected lasers had been operating continuously for more than a year at room temperature (typically 30°C). On the basis of the data obtained, the group was able to conclude that "continuous room-temperature operation of these devices as lasers with power outputs exceeding 1 mW per laser face for times in excess of 100,000 h is possible." This was an important feasibility demonstration. However, it served to reinforce the urgency of finding ways to confidently "accelerate" diode aging so that lasers tested for short periods could be installed in the field with the expectation that they would last for decades.



◀ Fig. 2. Schematic diagram of a proton-bombardment-delineated stripe geometry GaAs/GaAlAs semiconductor laser with a "window" on the substrate side. Note the four epitaxial layers of different composition grown by liquid-phase epitaxy [B. C. De Loach, Jr., B. W. Hakki, R. L. Hartman, and L. A. D'Asaro, Proc. IEEE **61**, 1042 (1973)].

By early 1977, with continued work on growth and process improvements, screening techniques, and protocols for accelerated aging, it was felt that, for a set of randomly selected lasers, it was possible to confidently predict a median lifetime at 22°C at 34 years and a mean time to failure at 22°C at 1.3 million hours (>100 years). The so-called "million hour paper," which was published in 1977 [3], demonstrated that it was possible to construct semiconductor laser devices with very long lifetimes.

Soon after these results were published, the author attended a conference in England on the general subject of light emission from semiconductors. During the Q&A, the head of the laser development program at the Standard Telecommunications Laboratory asked, publicly and rather pointedly, "Dick, would you please tell us the secret of your reliability success?" The author puzzled for a moment and then blurted into the microphone, "We do everything very carefully." This brought a good deal of laughter from the audience, but it was not intended as a joke. It took some years to convince skeptics that the success of the Bell group's development program required the solution of hundreds of problems, innovation by scores of outstanding well-motivated people, millions of dollars, systematic iteration, and a good deal of time. Perhaps its key achievement was the "proof of principle" that semiconductor lasers with long lifetimes were possible—a little like Roger Bannister's four-minute mile. In years since, it has appeared that most business and political leaders, as well as scientists who have not been involved in difficult high-tech development programs, do not appreciate what it takes to succeed with these types of endeavors.

In any case, after the group's considerable reliability achievements, the hard parts of the laser development program still lay ahead. The words of the great statesman Winston Churchill, referring to much more serious issues than ours, provided some encouragement: "Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning."

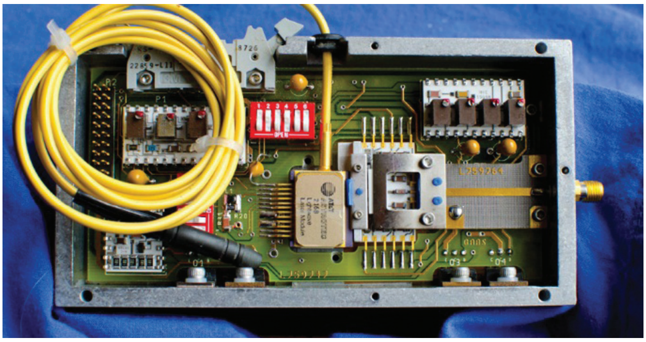
As the Bell group became better able to fabricate and age lasers, the testing of device characteristics and the ability to analytically model these devices were also refined. These developments greatly aided the early identification of lasers with deficiencies and also pointed the way to eliminating those problems.

The first applications of these lasers in the Bell System were in system experiments that were not intended to carry commercial traffic. These used 50-μm-core multimode fiber and data rates of 45 and 90 Mb/s. After that, they were tried in short-distance trials carrying live traffic, including a successful May 1977 installation in which fibers were used to connect three telephone central offices in downtown Chicago. The small physical size and large capacity of the fiber system helped to relieve crowding in the underground (and sometimes underwater) ducts that connected the offices. Then the trials became ambitious: In February 1980 at the Winter Olympics at Lake Placid, New York, the television feed was carried over an experimental optical fiber system and broadcast around the world. Fingers were crossed! In the end, it was fabulous to see the "Miracle on Ice" performance of the U.S. men's ice hockey team via the superior television picture made possible by our fiber system (see Fig. 3). These first-generation lasers were subsequently used in the fiber systems for the Northeast corridor and many other terrestrial trunk applications. Technology "proof of principle" had become "technology of choice!"

The second-generation lasers were designed for use in the 1.3-μm window of the improved, single-mode, 5-μm-core diameter fibers. Their buried-heterostructure design made use of two epitaxial growth sequences with an etching step in between. Figure 4 shows a 1.7-Gb/s transmitter developed by Optical Society Fellow Richard G. Smith and his group that made system implementation possible.

The complex fabrication process ultimately produced very high-yield, high-performance, high-reliability buried-heterostructure lasers that stayed in volume production from about 1984 to 1997. These multimode lasers could be used at data rates up to about 2 Gb/s. They were the mainstay of the Bell System's 417-Mb/s applications and, later, its FT series G 1.7-Gb/s applications providing the first high-speed 1.7-Gb interconnects among some 200 major U.S. cities. These and subsequent lasers came to possess such high reliability and could be applied in terrestrial trunk and undersea applications because the Bell System group became increasingly able to screen out lasers that had non-fundamental modes of degradation. Short-duration high-stress testing, specific to the individual laser design in the laser certification process, was used.

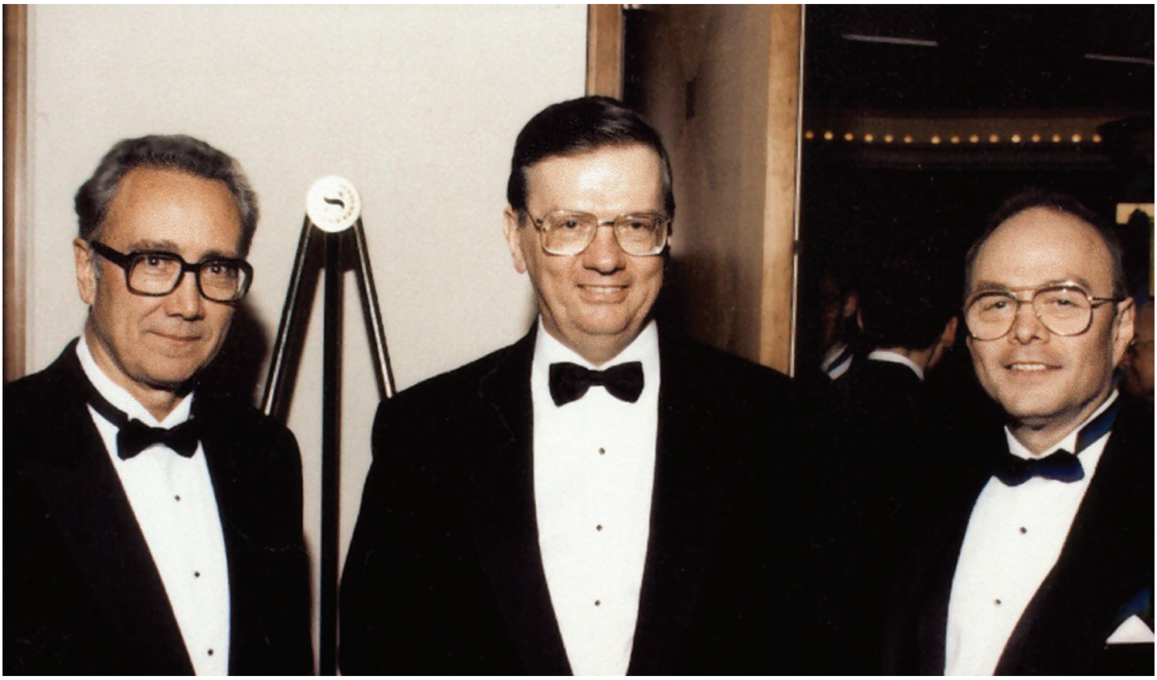
▲ **Fig. 3.** The televised feed for the 1980 Olympic Ice Hockey matches (like the one between Canada and the Netherlands shown here), including the famous "Miracle on Ice" game, was carried over an experimental fiber optic system.



▲ **Fig. 4.** 1.7-Gb/s transmitter.

Subsequently, more-sophisticated InP-based lasers, including designs with distributed feedback gratings to produce a single, stabilized wavelength [4] were designed, developed, and manufactured by the group. An electro-absorption modulator was later incorporated, on the same chip, into this design. Descendants of these devices—operating at data rates as high as 40 Gb/s, but more typically at 10 Gb/s—are useful for wavelength-division-multiplexing applications and remain in volume production in the United States, Japan, and other parts of the world. They make use of the "ultimate" low-loss 1.5–1.6-μm region in modern single-mode fibers. Metal organic chemical vapor deposition technology has now substantially replaced LPE in diode laser manufacture.

During these long, difficult years, the author sometimes pondered the meaning of Wolfgang Pauli's characterization of condensed-matter physics as "Schmutzphysik." Did he mean simply that it was complex and therefore hard? Did he mean that it was difficult literally because impurities (dirt) at unheard-of small concentrations affect everything? Or did he simply mean that any elegant physics involved was hidden in an opaque matrix of mud? At times, the author thought of the group's researchers as the "mudders." Fortunately, they ended up finding gold. Bob Rediker, a professor at MIT and MIT Lincoln Labs, expressed his view of the work leading to long-lived diode lasers as follows: "In the 1980s and early 1990s, I mounted a campaign with



▲ Fig. 5. B. C. DeLoach, R. W. Dixon, and R. L. Hartman receiving the IEEE Gold Medal for Engineering Excellence in 1993.

others to insist that those who by much hard work made inventions practical be honored. In particular, I wanted recognition for the team at Bell Telephone Laboratory. They had increased the mean time to failure at room temperature of the double-heterostructure GaAs-based laser from several minutes in 1970 to an extrapolated 8 million hours in 1978." Rediker's efforts along with others led to B. C. DeLoach, R. W. Dixon, and R. L. Hartman receiving the IEEE Gold Medal for Engineering Excellence in 1993 for this work (see Fig. 5).

The group's efforts in the 1970s, 1980s, and 1990s were aimed at Bell System applications in long-distance, high-volume voice, data and video transmission—both on land and undersea. Today, essentially all terrestrial and undersea telecommunications, data, and television traffic above the local distribution level is carried in fiber using lasers as sources. The Internet would not be possible without these laser devices. Undersea cables with long repeaterless spans (approaching 10,000 km) now often have the high-performance lasers that encode digital information only at the land ends. Much simpler continuously operating lasers, which carry no signal information, are used to pump fiber amplifiers that are periodically spaced under the sea. Data rates in a single fiber, using very-high-speed modulation and wavelength division multiplexing, in high-volume applications, can approach 1 Tb/s—20,000 times higher than the group's initial 45-Mb/s rates!

The program also supported what was then called "fiber-to-the-home," or colloquially "the last mile." This application took longer to become a reality because of the breakup of the Bell System and the high costs of serving individual customers. It was pleasing, and a little nostalgic, when about five years ago Verizon brought their laser-based FiOS product to the author's home. On the consumer products side, it has been extremely satisfying to witness the unexpectedly fast and widespread application of lasers in products such as printers and CD/DVD players and/or the dramatic price reductions made possible by these high-volume applications. Through the efforts of thousands of scientists and engineers throughout the world, both the programs the author worked on and their subsequent applications have succeeded beyond his wildest dreams.

The author is grateful to each one of the scores of professional scientists, technologists, and many others who contributed to the success of the Bell Laboratories semiconductor laser development program during the last decades of the twentieth century. It was fun being along for the ride.

References

1. I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumski, “Junction lasers which operate continuously at room temperature,” Appl. Phys. Lett. **17**, 109–111 (1970).
2. B. C. De Loach, Jr., B. W. Hakki, R. L. Hartman, and L. A. D’Asaro, “Degradation of CW GaAs double-heterojunction lasers at 300 K,” Proc. IEEE **61**, 1042–1044 (1973).
3. R. L. Hartman, N. E. Schumaker, and R. W. Dixon, “Continuously operated (Al, Ga)As double-heterostructure lasers with 70°C lifetimes as long as two years,” Appl. Phys. Lett. **31**, 756–759 (1977).
4. J. L. Zilko, L. Ketelsen, Y. Twu, D. P. Wilt, S. G. Napholtz, J. P. Blaha, K. E. Strege, V. G. Riggs, D. L. van Haren, S. Y. Leung, P. M. Nitzche, and J. A. Long, “Growth and characterization of high yield, reliable, high-power, high-speed, InP/InGaAsP capped mesa buried heterostructure distributed feedback (CMBH-DFB) lasers,” IEEE J. Quantum Electron. **25**, 2091–2095 (1989).
5. D. P. Wilt, J. Long, W. C. Dautremont-Smith, M. W. Focht, T. M. Shen, and R. L. Hartman, “Channelled-substrate buried-heterostructure InGaAsP/InP laser with semi-insulating OMVPE base structure and LPE regrowth,” Electron. Lett. **22**, 869–870 (1986).

Terabit-per-Second Fiber Optical Communication Becomes Practical

Guifang Li

Humans used optical signals intuitively for the purpose of communication in ancient times. Modern day optical communication systems are instead based on the fundamental understanding of information theory and technological advances in optical devices and components. The Optical Society (OSA) played a vital role in making fiber-optic communication practical for the information age.

It is well known that the capacity of a communication channel is constrained by the Shannon limit, $W \log_2(1 + S/N)$, where W is the spectral bandwidth and S/N is the signal-to-noise ratio (SNR). The bandwidth of a communication channel is proportional to the carrier frequency, which is on the order of 200 THz for visible or near-infrared light. Therefore, a small fractional bandwidth around the optical carrier can provide a capacity much larger than the limited capacity supported by the spectrum of radio-frequency (RF) waves or microwaves [1]. The SNR of a communication channel is proportional to the received power and inversely proportional to the noise and distortion. The invention of the laser, which can produce high-power coherent optical radiation at the transmitter, fueled the migration from RF/microwave communication to optical communication. In fact, the first patent on lasers (more precisely masers) by Nobel Laureates Charles Townes and Arthur Schawlow, both OSA Honorary Members, was entitled "Maser and maser communication systems."

To make optical communication practical, however, the received optical power (not only the transmitted power) must be much stronger than the noise. This requires a low loss optical transmission channel. The loss in free-space transmission is determined by diffraction, which is much larger than that of RF/microwave in appropriate cables. Fortunately, light can also be guided by total internal reflection, a phenomenon known since the mid-nineteenth century. An optical fiber with a high-index core surrounded by a lower-index cladding can support guided "modes" inside the dielectric cylindrical waveguide that propagate without experiencing radiative loss [2]. As a consequence, the loss of the optical fiber is dominated by material loss. Glass fibers were initially deemed impractical for communication systems, as the measured attenuation was >1000 dB/km.

In 1966, Kao and Hockham showed that the measured losses were due to impurities rather than fundamental loss mechanisms and, without impurities, glass fibers could achieve losses below 5 dB/km. They also identified that fused silica fiber could have the lowest losses. OSA Fellow Dr. Charles Kao was awarded the 2009 Nobel Prize in Physics for his "groundbreaking achievements concerning the transmission of light in fibers for optical communication," which has fundamentally transformed the way we live our daily lives. It is the invention of the silica optical fiber and the semiconductor laser with significantly long life that ushered in the era of modern optical communication. (These inventions are described in separate essays in this section of this book.)

The first-generation fiber-optic communication system in the 1980s used multimode fibers and 0.8-μm multimode Fabry–Perot semiconductor diode lasers, supporting a data rate

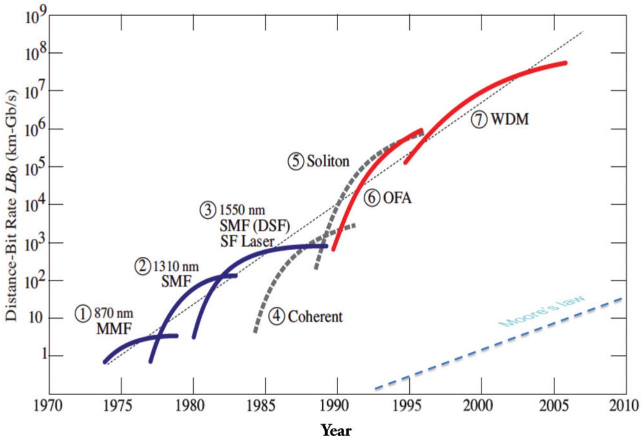


Fig. 1. History of fiber-optic communication systems. (Courtesy of Tingye Li, Alan Willner, and Herwing Kogelnik.)

of 45 Mbit/s [3], which was orders of magnitude larger than that of the microwave cable systems then in use. Since then, the capacity of optical fiber communication systems has grown in leaps and bounds. Throughout its history, fiber-optic communication has invented and reinvented itself many times over, as shown in Fig. 1, making terabits per second (Tb/s) practical. For example, the second-generation fiber-optic communication system operated at 1310 nm using single-mode fibers and single-mode semiconductor diode lasers. This brought about two improvements over the first-generation systems. First, the 0.3-dB/km loss of optical fiber at 1310 nm is much lower than 3 dB/km at 870 nm, which helped to overcome noise. Second, 1310 nm is the zero-dispersion wavelength for standard single-mode fiber. All of these different stages of technology development overcame different physical limitations of the optical communication system, pushing capacity toward the Tb/s fundamental limit. The physical limitations for fiber-optic communication arise from noise and distortion.

First let us focus on the sources of noises, which are closely related to modulation formats. Before 1980, the modulation format for optical communication systems was intensity-modulation direct detection, which is thermal noise limited with a sensitivity of thousands of photons/bit. In an effort to overcome thermal noise, the third-generation optical communication systems moved to 1550 nm, which is the minimum-loss wavelength for single-mode fibers, to increase the received optical power. As the additional power budget allowed gigabits-per-second transmission, distortions due to fiber dispersion could sometimes be the limiting factor. So in some third-generation systems, dispersion-shifted fibers for which the zero-dispersion wavelength was shifted to 1550 nm through proper design of the fiber index profile were used. In such systems, the capacity was still limited by thermal noise. Thus, starting from the mid-1980s, the optical communications community embarked on the development of coherent detection. Phase-shift keying (PSK) using coherent homodyne detection is limited by the shot noise of the local oscillator, and for binary PSK the sensitivity is 9 photons/bit, two orders of magnitude better than the thermal noise limit. However, coherent optical communication did not advance into commercial deployment because (1) phase locking and polarization management of the local oscillator was too complex and unreliable, and (2) the advent of the erbium-doped fiber amplifier (EDFA) made it unnecessary.

As early as 1964, rare-earth metal-doped glass fiber was proposed and demonstrated as a gain medium for optical amplification [4]. However, it was not until the late 1980s when two groups published work demonstrating high-gain EDFAs for fiber-optic communication—first by the group led by David Payne [5] and then by Emmanuel Desurvire [6]—that EDFA revolutionized the field of optical communication. Payne and Desurvire received the John Tyndall Award from OSA in 1991 and 2007, respectively. In terms of noise performance, optical pre-amplification (using an EDFA in front of the photodetector) changes the dominant noise source to the amplified spontaneous emission of the EDFA rather than the thermal noise of the photodetector. The fourth-generation optical communication system employed pre-amplified direct detection, which has a sensitivity of 39 photons/bit. (An essay on fiber optical amplifiers is in this section of this book.)

In fact, the gain bandwidth of an EDFA is ~3 THz, much wider than the single-channel bandwidth, which is limited by the speed of electronics. As a result, EDFAs enabled the fifth generation of wavelength-division-multiplexed (WDM) optical transmission systems. In these systems independent data streams are simultaneously transmitted on multiple wavelength channels in a single fiber and amplified together in a single EDFA, similar to frequency-division multiplexing in radio communication. WDM systems, championed by Dr. Tingye Li, 1995 OSA President, provided a multiplicative expansion of the fiber-optic bandwidth and thus multiplicative growth in fiber-optic communication system capacity. The development of WDM systems, a major leap forward in optical communication, began in the late 1990s.

Now let us focus on distortions in fiber-optic communication. Chromatic dispersion and polarization-mode dispersion (PMD) are linear distortions that exist in optical fibers. With the availability of EDFAs, optical power became an abundant resource that was extremely useful in combating the effects of noise. But high optical power also introduced nonlinear distortions in optical fiber that do not have analogies in radio communication. This is because optical fibers exhibit an intensity-dependent refractive index called the Kerr nonlinearity. Kerr nonlinearity leads to self-phase modulation and intensity-dependent spectral broadening, which in conjunction with dispersion ultimately leads to amplitude noise and timing jitter. In addition, for WDM systems Kerr nonlinearity also manifests itself in cross-phase modulation and four-wave mixing (FWM). Four-wave mixing requires phase matching of the four waves or momentum conservation of the four photons. As a result, FWM is very strong in dispersion-shifted fiber and can be effectively suppressed in fibers with a small amount of dispersion. For WDM systems, FMW is a dominant nonlinear distortion. Therefore, WDM systems must have dispersion to avoid strong nonlinearity. But dispersion is detrimental because it introduces linear distortion. The solution to this dilemma is the dispersion- and nonlinearity-managed WDM system consisting of fibers with positive dispersion and negative dispersion in cascade. And because local chromatic dispersion is never zero, nonlinear distortions are suppressed. As a result, the net overall chromatic dispersion is zero, so there is no linear distortion. Dispersion- and nonlinearity-managed WDM systems account for the majority of undersea systems all over the world. Dr. Andrew Chraplyvy and Robert Thack received the John Tyndall Award from the OSA in 2003 and 2008, respectively, for their contribution to the fundamental understanding of linear and nonlinear distortions.

After the turn of the new millennium, coherent optical communication made a comeback. This was made possible by advances in digital signal processing (DSP) and large-scale application-specific integrated circuits. In sixth-generation digital coherent optical communication, hardware phase locking and polarization management in conventional coherent optical communication of the 1980s were replaced by digital phase estimation and electronic polarization demultiplexing using multiple-input–multiple-output techniques. On the surface, it may seem incremental to migrate into coherent optical communication when the improvement in sensitivity is rather limited and the price to pay is the complicated DSP. The answer lies in the fact that DSP can perform not only phase and polarization management but also a number of other functionalities better than or impossible for optics in WDM systems. First, digital coherent communication enables electronic compensation of all linear distortions/impairments, including chromatic dispersion, PMD, and non-ideal frequency response of all components in the transmitter and receiver. Electronic dispersion compensation eliminates the need for dispersion-compensation fibers (DCFs), which leads to even less nonlinearity considering that DCFs have a small effective area and fewer amplifiers, and thus reduced noise. Reduction in both nonlinear distortions and noises improves system performance. Theoretically, it is even possible to use DSP to compensate nonlinear distortions. Digital coherent optical communication truly brought current fiber-optic systems to the fundamental capacity limit, the so-called nonlinear Shannon limit, of the single-mode fiber.

Fueled by emerging bandwidth-hungry applications and the increase in computer processing power that follows Moore's law, internet traffic has sustained exponential growth. This trend is expected to continue for the foreseeable future. As today's dense (D)WDM optical communication technology has already taken advantage of all degrees of freedom of a lightwave in a single-mode fiber, namely, frequency, polarization, amplitude, and phase, further multiplicative growth has to explore new degrees of freedom. Since the 2010 Optical Fiber Communications Conference, mode-division multiplexing in which every mode in a multimode fiber transmits independent information has emerged as a promising candidate for the next multiplicative capacity growth for optical communication. Suffice it to say that innovations for petabits-per-second (Pb/s) fiber-optic communication will continue in the foreseeable future.

References

1. R. Kompfner, "Optical communications," Science **150**, 149–155 (1965).
2. D. Hondros and P. Debye, "Elektromagnetische Wellen an dielektrischen Draehten," Ann. Phys. **32**, 465 (1910).
3. T. Li, "Advances in optical fiber communications: an historical perspective," IEEE J. Sel. Areas Communic. **1**, 356–372 (1983).
4. C. J. Koester and E. Snitzer, "Amplification in a fiber laser," Appl. Opt. **3**, 1182–1186 (1964).
5. R. J. Mears, L. Reekie, I. M. Jauncey, and D. N. Payne, "Low-noise erbium-doped fibre amplifier operating at 1.54 μm," Electron. Lett. **23**, 1026–1028 (1987).
6. E. Desurvire, J. R. Simpson, and P. C. Becker, "High-gain erbium-doped travelling-wave fiber amplifier," Opt. Lett. **12**, 888–890 (1987).

1975–1990

Applied Nonlinear Optics

G. H. C. New and J. W. Haus

The recent fiftieth anniversary celebrations marking the invention of the laser and the birth of modern nonlinear optics were major historical milestones. Theodore Maiman's observation of laser action in ruby in May 1960 [1] provided the essential tool that enabled Peter Franken's team at the University of Michigan to perform their legendary 1961 experiment in which they saw optical second harmonic generation for the first time [2]. From this small beginning, nonlinear optics has grown into the vast and vibrant field that it is today.

The Optical Society Centennial provides an opportunity to reflect on developments in nonlinear optics in the intervening years and, specifically, to focus on some of the highlights in the development of the field between 1975 and 1990. The theoretical foundations of optical frequency mixing were laid by Nicolaas Bloembergen's Harvard team in a seminal 1962 paper [3], which was prescient for introducing innovative ideas that strongly influenced later developments in the field; some specific examples will be mentioned later. In 1979, Nicolaas Bloembergen (see Fig. 1) was awarded The Optical Society's Ives Medal, the society's highest award. He won a quarter share of the 1981 Nobel Prize "*for his contribution to the development of laser spectroscopy*," in addition to his pioneering work on nonlinear optics.

By the early 1970s, many of the conceptual foundations of nonlinear optics had been laid, and a remarkable number of crude experimental demonstrations of techniques that are now routine had been performed. Progress over the ensuing decades was often prompted by advances in laser technology and, crucially, in materials fabrication. Suddenly it would become possible to implement an experiment so much more effectively than previously that it would soon become an established laboratory technique, or might even form the basis of a new commercial product.

A major achievement of the period was the fabrication of layered crystalline structures in which phase-matching is determined by the periodicity of the layers. Remarkably, this "quasi-phase-matching" (QPM) technique was originally suggested in the 1962 Harvard paper mentioned earlier [3], and it is a prime example of a principle that took more than two decades of gestation between original inspiration and final fruition.

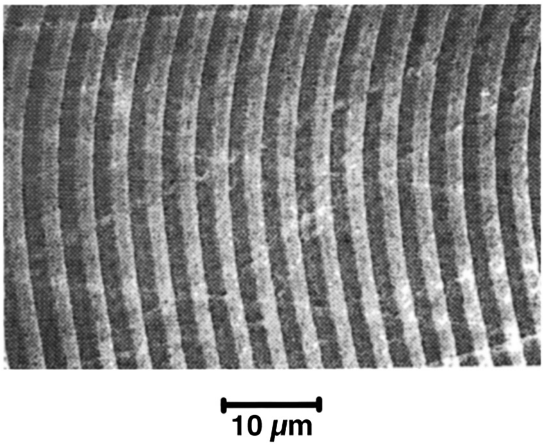
Quasi-phase-matching materials have periodically reversed domains, each one coherence length thick. The finished product is like a loaf of sliced bread in which alternate slices (of anisotropic crystal) are inverted (see Fig. 2). The problem is that each "slice" has to be only a few micrometers thick, so it would be a little thin for one's breakfast toast! It took more than two decades to develop the sophisticated crystal growth techniques needed to fabricate media with such thin layers. Today, QPM is routine; indeed many researchers have abandoned traditional birefringent phase-matching altogether. The most well-known QPM medium is perhaps periodically poled lithium niobate (abbreviated PPLN and pronounced "piplin"), and practical devices of high conversion efficiency are commercially available. In 1998, Robert Byer (see Fig. 1) and Martin Fejer were awarded The Optical Society's R. W. Wood Prize "*for seminal contributions to quasi-phase matching and its application to nonlinear optics*." More recently, in 2009, Robert Byer received the Ives Medal, The Optical Society's most prestigious award.

The need for tunable coherent light sources to replace tunable dye lasers drove the development of solid-state devices; these are not subject to messy chemical spills, and the tuning ranges achievable in a single medium can extend from the ultraviolet to the mid-wave infrared (3–5-μm) regimes.

▲ **Fig. 1.** Images of five scientists who have made major breakthroughs in the development of nonlinear optics. From top left to bottom right they are: Nicolaas Bloembergen, Robert L. Byer, Chuangtian Chen, Linn F. Mollenauer, and Stephen E. Harris. (Bloembergen, Mollenauer, and Harris photographs courtesy of AIP Emilio Segre Visual Archives, Physics Today Collection; Byer photograph courtesy of AIP Emilio Segre Visual Archives, Gallery of Member Society Presidents; Chen photograph courtesy of Professor Chen Chuangtian.)

An important nonlinear optical process for creating a wideband coherent light source is optical parametric generation. This is essentially sum-frequency generation (the generalized version of second harmonic generation) running in reverse. A high-frequency "pump" wave drives two waves of lower frequency, known as the "signal" and the "idler"; in photon language, the pump photon divides its energy between the signal and idler photons. Without a seed to define a particular frequency band, the signal and idler grow from noise, with frequencies determined by the phase-matching conditions. An optical parametric amplifier is a device of this kind with a signal or idler seed to fix the operating frequency. If the gain is high, the conversion efficiency can be quite large, even for a single-pass system. However, the efficiency can be greatly improved by placing the nonlinear medium within a well-designed cavity, creating an optical parametric oscillator (or OPO).

The first OPO was demonstrated by Giordmaine and Miller as early as 1965, but subsequent progress was slow, largely because nonlinear crystals of the necessary high quality were not available. Indeed, the OPO is another example of a device where technological capability lagged seriously behind concept. By the late 1980s, however, the introduction of new nonlinear materials coupled with progress in laser technology made it possible to realize low-threshold OPOs. Synchronous pumping can be employed, in which case the OPO is driven by a train of short pulses with the repetition rate matched to the round-trip time of the cavity. OPOs are now standard devices in the well-found laser lab.



▲ **Fig. 2.** SEM image of a periodically poled lithium niobate wafer. (Reproduced with permission from [4]. Copyright 1990, AIP Publishing LLC.)

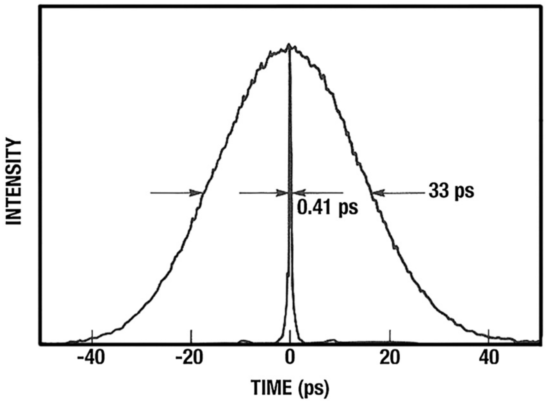
A number of new nonlinear materials that are now household names were developed in the 1980s. Using theoretical tools as a guide, C.-T. Chen (see Fig. 1) and co-workers discovered nonlinear materials such as BaB_2O_4 (beta barium borate, or BBO) and LiB_3O_5 (lithium borate, or LBO), both of which are widely used today. Other materials studied since that time include orientational-patterned III-V semiconductors, ZGP (zinc germanium phosphide) and DAST (4-dimethylamino-N-methyl-4-stilbazolium). Using a range of different nonlinear optical interactions, these have played an increasingly important role in extending the range of tunable coherent sources to the long-wave infrared (8–12 μm) and beyond to the terahertz regime. In recognition of the central role of materials technology, The Optical Society sponsored a 1988 conference entitled "Nonlinear Optical Properties of Materials," and key results were published in a special issue of the *Journal of The Optical Society of America B* [5].

The nonlinear interactions mentioned so far are all second order, which also means that they involve the interaction of three waves. Third-order processes lead to a wide range of four-wave phenomena, which include third harmonic generation, self-phase modulation via the optical Kerr effect (nonlinear refraction), optical phase conjugation, and optical bistability, to name just a few. They also form the basis of much of nonlinear spectroscopy, and quantum optical effects too.

Many important applications are based on nonlinear refraction. In combination with diffraction, it is the essential ingredient in the formation of spatial solitons, while with group velocity dispersion, it is crucial in the control of temporal pulse profiles. The 1970s and 1980s saw rapid progress in the understanding of optical pulse propagation and the development of nonlinear pulse compression techniques. Most of the techniques involve judicious combinations of self-phase modulation (SPM) and group velocity dispersion (GVD). Both of these processes cause a pulse to acquire a carrier frequency sweep (or "chirp"), but the overall effect depends on whether the two processes work with or against each other and whether they occur simultaneously or in succession. If they act simultaneously and in opposition, pulse propagation is governed by the nonlinear Schrödinger equation, which supports optical solitons.

In the early 1970s, Hasegawa and Tappert had suggested that optical fibers offered the ideal environment for solitons, but it was not until 1980 that Mollenauer, Stolen, and Gordon at what was then still Bell Telephone Labs actually observed optical soliton propagation in a fiber. Later, in 1988, Mollenauer and Smith demonstrated the transmission of 55-ps pulses over 400 km by supplying Raman gain at 42-km intervals. The possible use of solitons in optical communications was vigorously pursued in the 1990s but has rarely been implemented commercially. Nevertheless, research on solitons (both temporal and spatial) had a significant impact on nonlinear optics and indeed on laser technology as well. Linn Mollenauer (see Fig. 1) was awarded The Optical Society's Charles Hard Townes award in 1997 for his work on optical solitons and their applications to data transmission. Earlier, in 1982, he had received the R. W. Wood Prize for his work on color-center lasers, which played a vital role in early soliton experiments.

The race to achieve ever shorter optical pulses began on the day Maiman demonstrated the first laser and is likely to run for as long as laser research continues. Its hallmark has always been the strong and highly productive synergy between nonlinear optics and laser development. On the one hand, nonlinear interactions are strengthened by the high peak power of short laser pulses, but nonlinear optical processes are themselves exploited in advanced laser systems to promote the generation of shorter pulses.



▲ **Fig. 3.** Eighty times compression of a pulse. (Reproduced with permission from [6]. Copyright 1984, AIP Publishing LLC.)

▲ **Fig. 4.** Supercontinuum generation using a prism to disperse the colors in the pulse. (Image courtesy of [7]. © 2008 SPIE, image credit: E. Goulielmakis, reprint_permission@spie.org.)

The basic principle of pulse compression involves the application of SPM and GVD in opposition (as for solitons), but in succession rather than simultaneously. The idea, which originated in the late 1960s, is to start by imposing SPM to broaden the pulse spectrum and create the bandwidth required to support a shorter pulse. The wideband signal is then compressed by using a dispersive delay line, usually based on a pair of diffraction gratings in a Z-shaped configuration, which has a similar effect to that of negative GVD. An attractive option is to introduce the SPM in an optical fiber since, for non-trivial reasons, the simultaneous effect of SPM and *positive* GVD produces stretched profiles that are ideal for efficient compression.

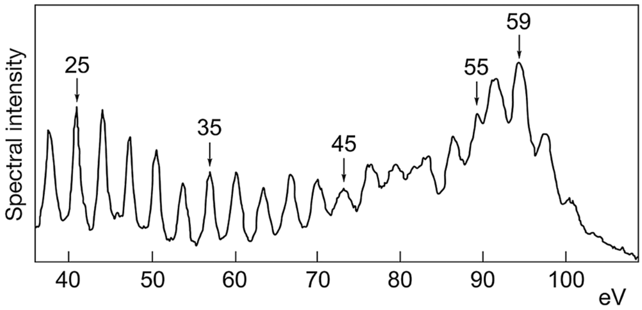
Fiber-grating compressors were first demonstrated in 1981, and there followed a series of record-breaking experiments that included the remarkable 1984 demonstration by Johnson, Stolen, and Simpson of a compression factor of ×80 (from 33 ps to 410 fs, Fig. 3). The culmination of this effort was the famous achievement of a 6-fs pulse by Fork, Shank, and Ippen in 1987, a result that held the world record for the shortest optical pulse for many years thereafter.

Other important nonlinear effects occur when pulses are launched in a fiber. Under suitable conditions, the combination of SPM and stimulated Raman scattering (SRS) creates a signal that extends over more than an octave in frequency bandwidth. A broadband signal of this kind is called a supercontinuum and has valuable applications in metrology and spectroscopy (Fig. 4).

Most developments in nonlinear optics in the 1960s involved solid media, especially crystals, although liquids also featured in experiments on the optical Kerr effect. By contrast, the 1970s and 1980s saw the beginning of work on the nonlinear optics of atoms and molecules in the gas phase that would come to full fruition in the 1990s and 2000s in effects such as high harmonic generation (Fig. 5), attosecond pulse generation, electromagnetically induced transparency, and slow light.

The early work on third harmonic generation in the inert gases in the late 1960s, and experiments on third, fifth, and seventh harmonic generation in metal vapors in the 1970s by Harris, Reintjes, and others, all exhibited characteristics typical of the perturbative (weak-field) regime, insofar as the conversion efficiency for higher harmonics fell away sharply. The first steps on the road that would later lead to the gateway into high harmonic generation were taken in the late 1980s. By that time, laser intensities of ~100 TW/cm^2 and above were becoming available, and some remarkable results on the inert gases were recorded that marked the entry into a new strong-field regime. For the lower harmonics (up to perhaps the ninth), the conversion efficiency dropped off as before, but higher harmonics lay on a plateau on which the efficiency remained essentially constant up to a well-defined high-frequency limit. The cut-off point could be extended further into the UV by increasing the laser intensity, although a saturation intensity existed beyond which no further extension was possible. These experiments laid the foundation for work in the following decade in which harmonics in the hundreds and even the thousands were generated.

An equally dramatic line of development involved atomic systems in which the main action involved three levels linked by two separate laser fields. A number of different effects of this kind were beginning to be studied in early 1980s, most of which exploited the effect of quantum interference in one way or another. Early examples included coherent population trapping and laser-induced continuum structure, both of which were prefigured to some degree in the much earlier work of Fano and others on Fano interference.



▲ Fig. 5. Experimental manifestation of high harmonic generation. (Courtesy of [8]. Copyright 2011, Cambridge University Press.)

In the mid to late 1980s, the effect of lasing without inversion (LWI) caused a particular stir, probably because it contradicted a principle that most people regarded as fundamental to laser physics, namely, that population inversion was an essential prerequisite of laser action. The scheme for LWI envisaged by Harris involved three levels in a pattern roughly resembling an inverted V, or a capital Greek lambda Λ. Under normal circumstances, laser amplification on one arm of the Λ would occur only if a population inversion existed between the two levels. Crucially, however, this restriction is removed if a strong laser field is tuned to the resonance frequency of the other arm of the Λ.

The simplest explanation of how LWI works involves another quantum interference process, highlighted by Harris in 1990, called electromagnetically induced transparency (EIT). A straightforward density matrix calculation shows that the absorption and dispersion characteristics of one of the transitions of the Λ are dramatically altered in the presence of the strong coupling field tuned to the other, and indeed that the absorption goes to zero on exact resonance. Quantum interference has in effect canceled out the absorption process that normally competes with stimulated emission, thereby enabling lasing to occur in the absence of a population inversion.

Stephen Harris (See Fig. 1) received the Ives Medal in 1991 for his pioneering work in nonlinear optics. The citation specifically mentioned his work on LWI and EIT.

Given the strict word limit that we have worked within, we have naturally been forced to be highly selective in choosing the topics to cover. Literally thousands of research papers on nonlinear optics presenting the work of many hundreds of researchers were written within the time frame covered in this chapter. In view of these numbers, it is inevitable that many people will consider topics we have left out to be more important than those we have included. We extend our apologies to the majority whose work it has not been possible to mention here.

References

1. T. Maiman, “Stimulated optical radiation in ruby,” Nature **187**, 493–494 (1960).
2. P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, “Generation of optical harmonics,” Phys. Rev. Lett. **7**, 118–119 (1961).
3. J. A. Armstrong, N. Bloembergen, J. Ducuing, and P. S. Pershan, “Interactions between light waves in a nonlinear dielectric,” Phys. Rev. **127**, 1918–1939 (1962).
4. G. A. Magel, M. M. Fejer, and R. L. Byer, “Quasi-phase-matched second-harmonic generation of blue light in periodically poled $LiNbO_3$,” Appl. Phys. Lett. **56**, 108–110 (1990).
5. C. M. Bowden and J. W. Haus, eds., “Nonlinear optical properties of materials,” J. Opt. Soc. Am. B **6** (April 1989).
6. A. M. Johnson, R. H. Stolen, and W. M. Simpson, “80× single-stage compression of frequency doubled Nd:yttrium aluminum garnet laser pulses,” Appl. Phys. Lett. **44**, 729–731 (1984).
7. J. Hewett, “Ultrashort pulses create ultrabroad source,” historical archive, Optics.org.
8. J. W. G. Tisch, Imperial College Attosecond Laboratory; reproduced from G. H. C. New, “Introduction to nonlinear optics,” Cambridge University Press 2011, with permission.

Linear and Nonlinear Laser Spectroscopy

M. Bass and S. C. Rand

Spectroscopy has been a fundamental part of optics ever since Newton first showed that white light could be dispersed into its constituent colors and later when Young showed that light was wavelike and provided a grating with which to measure its wavelength. The role of The Optical Society (OSA) in spectroscopy during the pre-laser era is described in an essay entitled "Spectroscopy from 1916 to 1940" in an earlier part of this book. The first experimental demonstration of a laser, a ruby laser, was made by Theodore Maiman in 1960, and soon after, in 1964, a Nobel Prize was awarded for prior theory on the topic to Charles Townes, Nikolay Basov, and Alexander Prokhorov. Additionally, parametric nonlinear optics was discovered by Peter Franken in 1961. The combination of lasers and nonlinear optics made possible incredible advances in spectroscopy leading to linear and nonlinear laser spectroscopy. Developments in this field were so numerous that this short account can only hope to capture the principal events of an important chapter in optics and OSA history.

Almost immediately upon the invention of the laser, scientists recognized that the two most obvious features of laser light, its high intensity and its spectral purity, were far beyond anything that had been available before. In less than a year following Maiman's ruby laser, Franken took advantage of its high intensity to demonstrate optical second harmonic generation and open up the field of nonlinear optics. This would lead to numerous nonlinear spectroscopies mentioned below. Different designs also permitted wide-ranging variations in the type of output obtainable from lasers. Very pure single-frequency light was created with continuous-wave lasers and very broad, supercontinuum sources were created with ultrashort pulse lasers. The availability of lasers with large or small bandwidths and short or long pulse durations enabled the development of dozens of new and powerful approaches to precision optical measurements.

The Debut of Laser Spectroscopy

In 1960 the extraordinarily high intensity and short pulse duration available from the first ruby lasers ushered in a whole new era of experimentation in optical spectroscopy. The shift to laser methodology was rapid. Consider that G. Dieke and H. Crosswhite published a landmark paper in 1963 on the spectroscopy of doubly and triply ionized rare earths. For emission experiments they used pulsed discharges with currents in excess of kiloamperes together with photographic emulsions. For absorption measurements they employed high-pressure mercury and xenon lamps. Yet Dieke's student, S. Porto, who had labored to record infrared spectra of molecular hydrogen with the same apparatus only a few years earlier, was at that very moment pioneering the use of lasers in revolutionary spectroscopic techniques at Bell Labs in Murray Hill. There, Porto and his colleagues made the first observations of scattering from F-centers and spin waves, and introduced resonant Raman laser spectroscopy for the study of solids. Porto was a Fellow of OSA, and when he returned to Campinas, Brazil, in 1974 he was also elected a Fellow of the Brazilian Academy of Science. The seeds of a quiet revolution in optics had been sown as far away as Brazil. This can be considered a key starting point in the internationalization of OSA as it heralded widespread scientific exchange between the United States and many other countries.

Time-domain laser spectroscopy offered optical measurement capabilities on time scales that were six orders of magnitude faster than stroboscopes. Pump–probe experiments with picosecond pulses could time-resolve the fastest luminescent processes and follow the pathways of rapid chemical reactions. Dynamic grating spectroscopies soon lent sophistication to the dynamical processes that could be read out from the interference patterns formed by intersecting beams in various systems. Processes that produced no luminescence at all, such as energy transport among excited states in molecular crystals (coherent exciton migration), began to be investigated using transient grating approaches.

The realization that all systems possessed finite third-order susceptibilities and could easily be phase-matched to yield intense signals led to widespread popularity of coherent four-wave-mixing spectroscopy. Degenerate four-wave mixing in a counterpropagating pump geometry came into vogue. Another approach was coherent anti-Stokes Raman (CARS) spectroscopy devised by P. Maker and R. Terhune. This and other "coherent spectroscopies" not only achieved high resolution but gave signal waves that conveniently emerged from the sample as beams. As a consequence they are still used today to study molecular dynamics in chemistry.

Monochromaticity, wavelength control, and frequency stabilization improved steadily throughout the late 1960s. Barger and Hall reported a versatile frequency-offset locking technique in 1969 that permitted the frequency of one laser to be tuned relative to that of a second laser locked to a saturated absorption feature of methane that was a candidate for an absolute frequency reference. Their experiment demonstrated tunable control over the frequency of light to a precision of ~1 kHz for periods as long as an hour. For the first time this hinted at the possibility of frequency references and clocks based on optical schemes rather than radio frequency sources.

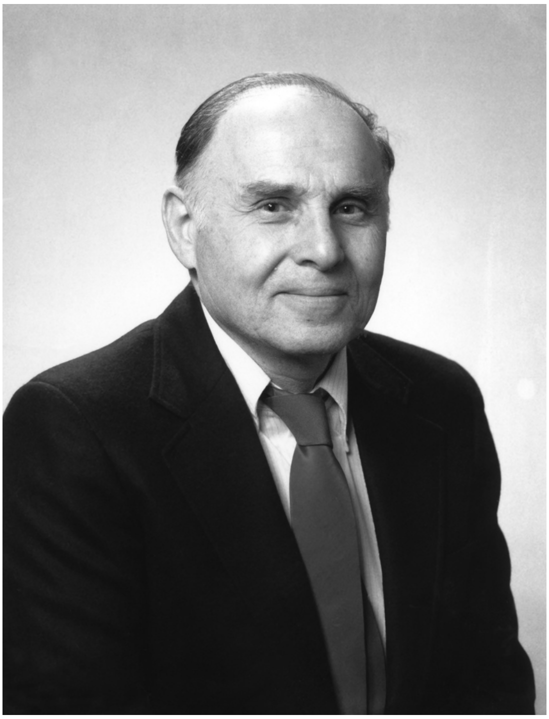
Optical modulation spectroscopies yielded still other measurement tools. When more than one transition of an atom was excited by a coherent optical pulse, excited-state fine or hyperfine structure produced modulation effects in the emission known as "quantum beats." At Columbia, D. Grischkowsky and S. Hartmann extracted frequency-domain splittings from time-domain photon echo signals in rare-earth-doped solids by simply Fourier transforming their data. This resolved the excited-state hyperfine structure with sub-megahertz precision and provided a beautiful example of the reciprocity between time- and frequency-domain measurements. In atomic spectroscopy the method of quantum beats also proved to be effective in resolving extremely fine splittings of energy levels in atomic vapors.

Gradual improvements in laser frequency control and methods of locking lasers together had the effect of encouraging researchers to think that the use of more than one laser in an experiment might eventually become possible, or even routine. The idea still seemed futuristic in 1972, so it came as quite a shock when the speed of light was redefined that year in a remarkable experiment by K. Evenson and his colleagues, who determined the speed of light to ten significant figures with an entire room full of frequency-locked lasers.

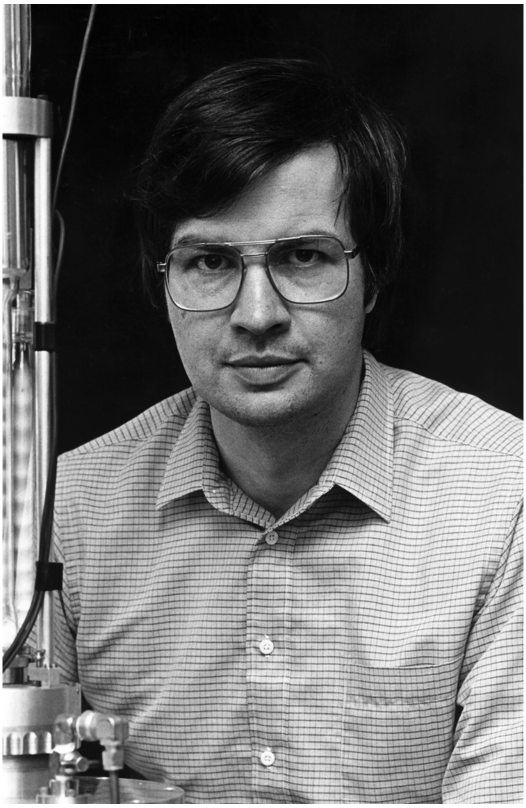
Following this, H. Dehmelt trapped ions in free space at the University of Washington, a feat for which he and W. Paul would share the 1989 Nobel Prize in Physics. A. Ashkin (see Fig. 1) at Bell Labs, and P. Toschek and H. Walther in Germany were thinking of ways to trap and cool individual neutral atoms. W. E. Moerner reported that single, isolated centers could be interrogated spectroscopically even in the complex environment of solids. The field of spectroscopy was poised to take on the challenges of laser cooling, Bose–Einstein condensation (BEC), single-molecule spectroscopy, and the control of trapped atoms for quantum information science.

Nonlinear Optics and Nonlinear Spectroscopy

A year after the (future OSA president) Peter Franken announced the experimental discovery of nonlinear optics at the University of Michigan in 1961, M. Bass observed sum frequency generation and then optical rectification. OSA meetings buzzed with the anticipation of additional possible discoveries of nonlinear phenomena. A general analysis of nonlinear interactions was published in September of 1962 by J. A. Armstrong and his colleagues. It indicated that an enormous number of nonlinear effects were possible at high laser intensities, and reports of experiments by other groups began to pour in. Nonlinear optics provided spectroscopists with



▲ **Fig. 1.** Arthur Ashkin. (AIP Emilio Segre Visual Archives, Physics Today Collection.)



▲ **Fig. 2.** Theodor Hänsch. (© OSA. Photo courtesy of Dr. W. John Tomlinson, Princeton, New Jersey.)

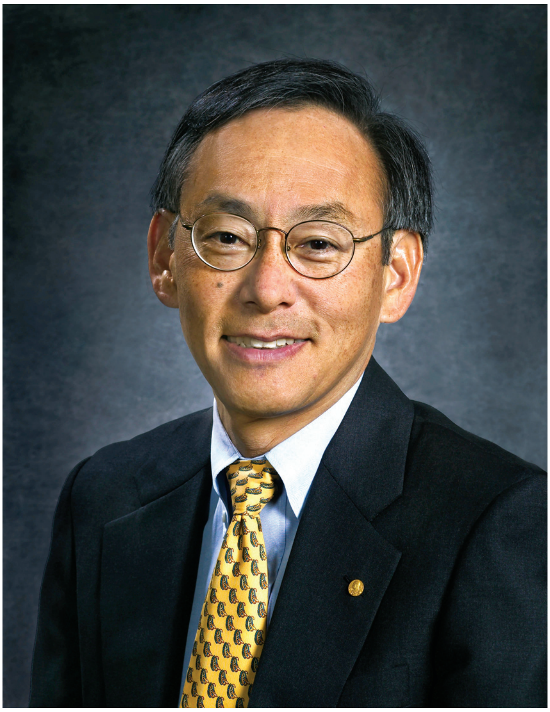
tools to reach otherwise inaccessible wavelengths, inaccessible spectral resolution, and unimagined short pulse durations.

The push for better resolution took a leap forward with the introduction of "Doppler-free" laser spectroscopy. C. Borde, T. W. Hänsch, A. L. Schawlow, V. Chebotayev, and V. Letokhov moved forward quickly to investigate its implications in Paris, Stanford, and Novosibirsk. It was widely recognized that spectral broadening due to motion of the atoms in a gas could be eliminated using a variety of methods: saturation spectroscopy, or 2-photon absorption, or by trapping atoms. The anticipated improvement in resolution from $\sim 10^4$ to $\sim 10^{11}$ using relatively simple experimental techniques was substantial enough that optical Lamb shift measurements could provide stringent tests of quantum electrodynamics. By 1975, research at Stanford based on 2-photon Doppler-free spectroscopy of hydrogen yielded a determination of the 1S Lamb shift for the first time. A concerted effort began to improve measurements of the Rydberg constant. At the time, the Rydberg constant was one of the most poorly determined fundamental quantities. In the decades that followed, its precision would improve a millionfold.

In 1977 the next tool for precision spectroscopy was introduced when the Ramsey fringe method was adapted for high resolution optical spectroscopy in Russia and in the U.S. This succeeded in extending the separated field technique from microwave to optical frequencies, for which Norman Ramsey received the 1989 Nobel Prize.

T. Hänsch (see Fig. 2) and A. Schawlow proposed a technique to stop atoms in order to improve spectroscopic resolution using laser radiation tuned below resonance. Their 1975 paper galvanized the spectroscopic community focused on precise frequency measurements. That same year laser spectroscopy on trapped barium ions was proposed, and by 1980 collaboration between Dehmelt and Toschek had succeeded in trapping a single Ba^+ ion in a quadrapole trap, cooling it to 10 mK with light, and observing its resonance fluorescence. Doppler-free spectroscopy of single Ba and Mg ions was on the horizon, and "optical clock" transitions became a topic of discussion. In 1982 H. Metcalf cooled a beam of neutral sodium atoms with a Zeeman "slower," and the next year D. Pritchard suggested a magnetic geometry to trap atoms. In 1985 S. Chu (see Fig. 3) reported an all-optical trap dubbed "optical molasses" and jointly with the MIT group announced an efficient magneto-optic trap in 1987 that could rapidly cool a variety of atoms to milli-Kelvin temperatures. Then in 1988 P. Lett of W. Phillips's group at NIST demonstrated cooling below the Doppler limit in alkali vapors. J. Dalibard and C. Cohen-Tannoudji (see Fig. 4) at ENS explained Lett's mechanism in a widely read 1989 publication in the *Journal of The Optical Society of America B*. Halfway around the world, researchers in Japan were in close pursuit, applying these advances to laser cooling of noble gases.

In 1995 these activities, originally motivated to improve spectroscopic resolution, culminated in the creation of a new form of matter. E. Cornell and C. Wieman observed BEC of Rb atoms at JILA in Colorado. By this time A. Schawlow and N. Bloembergen had shared the 1981 Nobel Prize for advances in spectroscopy. Chu, Cohen-Tannoudji, and Phillips were due to share this honor in 1997 for laser cooling. For producing and studying properties of BECs, Wieman, Cornell, and Ketterle would receive the Prize in 2001. J. Hall (see Fig. 5) and T. Hänsch would earn the Nobel prize in 2005 for the development of frequency "combs" that enabled tests of the variation of the gravitational constant and frequency references with uncertainties at the level of a few parts in 10^{15}.



▲ **Fig. 3.** Steven Chu. (Courtesy of U.S. Department of Energy.)

Laser Spectroscopy: An Enabling Science

The transition from spectroscopic research in the period 1960–2000 to its many applications had a long gestation period. D. Auston disclosed a method of generating single cycles of terahertz radiation in the 1980s. However, applications such as imaging through plastics and ceramics with terahertz waves would not become routine until the beginning of the twenty-first century. Similarly, as early as 1980, T. Heinz and Y. R. Shen found that second harmonic generation was allowed on the surfaces of centro-symmetric media but forbidden in their interior. IBM exploited this interaction to inspect silicon wafers for electronic circuits, but decades passed before species-specific structural and dynamic studies became popular with chemists. By the 1990s, experiments in the research groups of S. Harris and B. P. Stoicheff had established that opaque materials could be rendered transparent through quantum interference. This had immediate impact on spectroscopy and the generation of short wavelength

▲ **Fig. 4.** Claude Cohen-Tannoudji. (Photograph by Studio Claude Despoisse, Paris, courtesy AIP Emilio Segre Visual Archives, Physics Today Collection.)

radiation via nonlinear mixing. Yet once again a 20-year interval would pass before Rohlsberger was to achieve electromagnetically induced transparency at x-ray wavelengths, thereby hinting at the prospect of nuclear quantum optics.

There are other striking examples of how technological outgrowths of the last 50 years of spectroscopy continue to enable new science topics. The sub-Doppler laser cooling techniques of 1986 became tools for the fledgling field of quantum information. Only recently have they been applied to demonstrate 14-qubit entanglement with Ca^+ ions. Despite the frenzied activity in laser cooling and trapping that accompanied the race to achieve BEC, a quarter of a century also passed between the invention of "optical tweezers" by Ashkin for trapping particles and single cells and the studies of single biomolecules by S. Chu and others.

▲ **Fig. 5.** John Hall. (Courtesy of AIP Emilio Segre Visual Archives, Physics Today Collection.)

The Future

Over its 100-year lifespan, The Optical Society has been led by many accomplished scientists, many of whom were spectroscopists. It is partly for this reason that the society has been able to maintain a prominent role throughout an explosive period of scientific history that relied on precise spectral tests of new theories. Spectroscopists contributed to but also benefited from and were nurtured by the emphasis on fundamental science and the open, relaxed style of the Society, where many disciplines intersect. The vibrancy of OSA has rested on personal relationships fostered by the Society across ideological boundaries. OSA has followed a tradition of internationalization that began long before globalization made it necessary. Past president Art Schawlow understood how important international connections were for spectroscopy and science in general. He knew that when it came time for visitors from China, New Zealand, Canada, and Ireland to return home, they would inevitably take home part of his magic recipe for having fun with great science. They had learned that "You don't need to know everything to do good research. You just have to know one thing that isn't known," and of course you also had to be a spectroscopist! By sharing this attitude, Art was a great ambassador for the field of spectroscopy and for OSA itself. The rich history of both, and his encouraging message, accumulated in the hearts of his students and visitors. Current and future OSA members will sustain the unique strengths of the Society that account for its remarkable spectroscopic legacy and its future contributions.

Acknowledgement

Photos were provided by S. Svanberg, J. Hecht, H. van Driel, and the OSA archives. The authors wish to thank J. Eberley for a critical review.

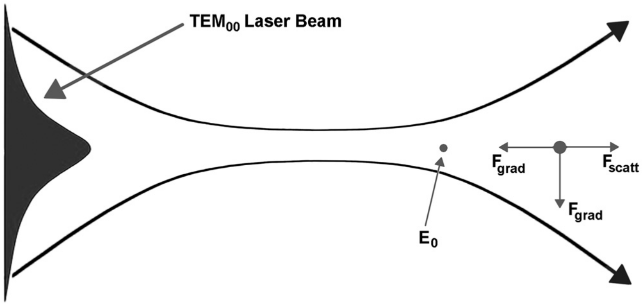
Optical Trapping and Manipulation of Small Particles by Laser Light Pressure

Arthur Ashkin

The invention of the laser has made possible the use of radiation pressure to optically trap and manipulate small particles. The particles can range in size from tens of micrometers to individual atoms and molecules. Laser radiation pressure has also been used to cool atoms to exceptionally low temperatures, enabling a new branch of atomic physics. See [1] for an extensive summary of the many varieties of work done with laser radiation pressure.

Inspired by a long interest in radiation pressure, in 1969 the author focused a TEM_{00} mode laser beam of about 30-μm diameter on a 20-μm transparent dielectric latex particle suspended in water. Strong motion in the direction of the incident light was observed. If the particle was off axis, at the edge of the beam, a strong gradient force component to the light force pulling the particle into the high-intensity region on the axis was observed. The particle motion was closely described by these two force components: one called the "scattering force" in the direction of the incident light and the other the "gradient force" in the direction of the intensity gradient. With these two components, and using two oppositely directed beams of equal intensity, it was possible to devise a stable three-dimensional all-optical trap for confining small particles. Particles moving about by Brownian motion that entered the fringes of the beam were drawn into the beams, moved to the equilibrium point, and were stably trapped. If the axial gradient force is made to exceed the scattering force, and this can be done, then a single-beam trap is possible, as shown in Fig. 1.

Because this was the first example of stable optical trapping, this discovery was submitted to *Physical Review Letters*. Since single atoms are just small neutral particles and should behave much as single dielectric spheres, it was postulated that trapping of single atoms and molecules should also be possible. At Bell Labs, if one wanted to submit a paper to *Physical Review Letters* one had to pass an internal review by the prestigious theoretical physics department to preserve the Lab's good name. So the author submitted a manuscript and it was rejected. Upon the recommendation of his boss, Rudi Kompfner, the inventor of the traveling-wave tube, the paper was resubmitted and was accepted with no problem [2]. A second theoretical paper was submitted to *Physical Review Letters* in 1970 on acceleration, deceleration, and deflection of atomic beams by resonance radiation pressure [3]. This was followed by a number of experiments on optical traps for micrometer-size solid spheres or liquid drops demonstrating optical levitation against gravity in air and as a function of pressure down to high vacuum and for various beam convergence angles. By using optical levitation in conjunction with feedback stabilization of the levitated particle's position, it was possible to study the wavelength dependence of the optical levitation forces with dye lasers. A series of complex size-dependent resonances were observed that were found to be in close agreement with Mie–Debye electromagnetic theory calculations. These results are probably the most exact confirmation of Maxwell's theory for light scattering by transparent dielectric spheres. The frequencies of these resonances allow one to determine the particle size and index of refraction to six or seven significant figures. Using the position stabilization technique it was possible to perform a modern



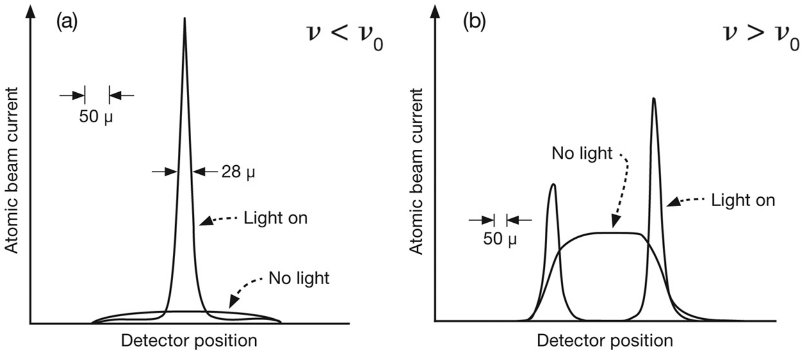
▲ Fig. 1. A single-beam optical trap for a high-index, transparent sphere. The laser beam is tightly focused such that the axial component of the gradient force exceeds the scattering force. E_0 is the equilibrium point at which the sphere is trapped.

version of the Millikan oil drop experiment for accurately determining the electric charge of a single electron.

Optical trapping of atomic vapors in high vacuum is more difficult than trapping macroscopic particles. One needs some form of damping for filling and holding atoms in an optical trap. Work was started in the early 1970s on accelerating, decelerating, and deflecting atoms with applications such as velocity sorting and isotope separation. T. Hänsch and A. Schawlow wrote an important early paper on optical cooling of atoms using the Doppler shift in a six-beam geometry for use in precision spectroscopy. They did not consider the possibility of optical trapping. In Russia, V. S. Letokhov and V. G. Minogen did experiments trying to stop sodium beams with chirped counterpropagating light beams, but failed. They were intending to trap atoms in a trap tuned a half-linewidth below resonance where cooling is a maximum. W. D. Phillips and H. W. Metcalf, inspired by Ashkin's first paper about atoms, also started work on atom slowing. They soon realized that the slowing difficulties experienced by Letokhov and Minogen were due to optical pumping, and in 1982 they successfully used a beam-slowing method based on a tapered magnetic field to completely stop the beam at a final temperature of about 0.1 K.

In 1978 Bjorkholm, Freeman, and the author carried out an experiment using tuning far from resonance that demonstrated dramatic focusing and defocusing of an atomic beam caused by the optical gradient forces [4] (Fig. 2). These striking results suggested that atom trapping would be possible if proper cooling could be achieved. It was realized that optical heating of atoms was a problem in achieving stable traps for cold atoms even for optimal tuning at a half-linewidth below resonance, where the cooling rate is a maximum, due to saturation. However, it was shown that deep trapping potentials were possible for two-beam traps and one-beam traps by tuning far-off resonance where saturation is greatly reduced. Two papers by Ashkin and Gordon addressed the details of laser cooling and heating and showed various ways of achieving adequate Doppler cooling.

In 1983 Steve Chu was transferred to our Holmdel Lab from the Murray Hill Lab. He was an experienced atomic physicist, but he did not know much about trapping at the time. He became interested and decided to join John Bjorkholm and the author in an attempt to trap atoms using lasers. This was at a time when we had some new bosses who decided that atom trapping would not work, and they tried unsuccessfully to discourage Bjorkholm and Chu from working with the author on this project. In spite of this pressure, Chu was given a quick lesson in atom trapping, and an effort was made to demonstrate the first optical trap for atoms. The first experiment was aimed at creating a collection of



◀ Fig. 2. Experimental demonstration of the focusing and defocusing of an atomic beam caused by the optical gradient force. (a) Laser tuned below resonance; atoms attracted to high intensity regions. (b) Laser tuned above resonance; atoms repelled from high intensity regions. (Redrawn from A. Ashkin, [IEEE J. Sel. Topics Quantum Electron. **6**(6), Nov./Dec., 2000.])

very cold atoms capable of being confined in the shallow atom traps. The experiment was based on the theoretical ideas proposed by Hänsch and Schawlow, mentioned earlier, and it worked beautifully. It provided a cloud of atoms having a temperature of about 240 μK, as expected, which is ideal for trapping. That cooling technique has become to be known as "optical molasses." Now that it was possible to generate cold atoms, Bjorkholm suggested trying a single-beam gradient trap in spite of its small size. The trap worked flawlessly, and shortly afterward, in December 1986, the work was featured on the front page of the Sunday *New York Times*. Surprisingly, a new trapping proposal by Dave Pritchard of MIT appeared in the same issue of *Physical Review Letters* as our trapping paper. It was for a large-volume magneto-optic scattering-force trap rendered stable via a quadrupole Zeeman-shifting magnetic field. The magneto-optical trap (MOT) is a relatively deep trap and is easily filled because of its large size. As later shown, it did not even require any atomic beam slowing.

Shortly after the atom trapping experiment, Chu left Bell Labs for Stanford and continued his atom trapping work. At Bell Labs, Bjorkholm and Ashkin turned to other work. Use of MOT traps dominated over dipole traps for atom work for about the next ten years. In 1997 the Nobel Prize in Physics was awarded to Chu, Phillips, and Cohen-Tannoudji for cooling and trapping of atoms.

In the lab, with the help of Joe Dziedzic, the author started looking at the use of focused laser beams as tweezers for the trapping and manipulation of Rayleigh particles. They made a surprising discovery one morning when while examining a sample that had been kept in solution overnight. Wild scattering was seen emanating from the focus of the trap. A joke was made about having caught some bugs. On closer examination it turned out that that this had happened. Bacteria had contaminated the sample, and they had fallen into the trap. The sample was placed under a microscope where the trapping could be observed in detail. In fact, the trap could be maneuvered to chase, capture, and release fast-swimming bacteria with green argon-ion laser light. If the laser power was turned up, "opticution" was observed; that is, the cell exploded. It was found that infrared YAG laser power was very much less damaging. Samples of E. coli bacteria obtained from Tets Yamane of Murray Hill were seen to reproduce right in the trap. Internal-surgery was performed in which the location of organelles was rearranged and the organelles were attached in new locations. The visco-elasticity of living cell's cytoplasm and the elasticity of internal membranes were also studied. This early work was the start of a new, unexpected, and very important application of laser trapping. A Nobel Prize winner at Bell Labs mentioned, amusingly in retrospect, that the author "should not exaggerate" by predicting that trapping would someday be important for the biological sciences.

Meanwhile, work to better understand optical molasses cooling of atoms was carried out at NIST and Stanford. Importantly, at NIST Phillips had made the surprising discovery of cooling to temperatures as low as 40 μK in "optical molasses." This was of great interest to those racing to achieve Bose–Einstein condensation (BEC) at very low temperatures and high densities. Anderson *et al.* won this race in 1995 using evaporative cooling from a magnetic trap reaching a temperature of about 170 nK at a density of $2{\times}10^{12}$ atoms/cm^3 with a loss of evaporated atoms by a factor of 500 from an original 10^7 atoms. Eric Cornell, Carl Weiman, and Wolfgang Ketterle received the 2001 Nobel Prize in Physics for the experimental demonstration of BEC.

The Nobel Committee in their 1997 press releases "Addendum B" on additional material mainly for physicists says "To become really useful one needed a trap deeper than the focused laser beam trap proposed by Letokhov and Ashkin and realized by Chu and coworkers in optical molasses experiments." On the contrary, far-off-resonance traps built according to Ashkin's design are the traps used in virtually every current Bose–Einstein experiment.

The story of the application of tweezer traps to biophysics and the biological sciences is more straightforward [5–7]. After the early work of the author on living cells, Ashkin and collaborators and Steven Block with Howard Berg showed the usefulness of optical tweezers for studying single motor molecules such as dynein, kinesin, and rotary flagella motors. Block and his co-workers continue to extend tweezer techniques to DNA replication and protein folding at even higher resolution (fractions of an angstrom) and lower force levels using super-steady optically levitated low-noise traps held in a helium gas environment.

Light-pressure forces are probably the smallest controllable and measurable forces in nature. Other low-force techniques such as atomic force microscopy (AFM) have their unique features but cannot function deep inside living cells, for example. Looking to the future, one expects the interesting work on motors and protein folding to continue. Perhaps we will see optical tweezers serving as gravitational wave detectors. Large improvements in atomic clocks have been made in the past using atomic fountain techniques. Recently another breakthrough has been made using ultracold optical lattice clocks approaching a stability of one part in 10^{18}. This achievement in time keeping by NIST has many potential applications.

The study of light is fundamental to physics. As such, one expects that applications of optical trapping and manipulation of particles by laser light pressure will continue well into the future.

The importance of using lasers for the trapping and cooling of atoms has been recognized by a number of prizes and awards, including the Nobel Prizes mentioned above. In addition, Arthur Ashkin has been recognized for his work in that field by The Optical Society (OSA) with the Charles H. Townes award in 1988, with the Ives Medal/Quinn Award in 1998, and by being elected an Honorary Member of OSA in 2010.

Many thanks to John Bjorkholm for his help in editing this essay.

References

1. A. Ashkin, Optical Trapping and Manipulation of Neutral Particles Using Lasers: A Reprint Volume with Commentaries (World Scientific, 2006).
2. A. Ashkin, “Acceleration and trapping of particles by radiation pressure,” Phys. Rev. Lett. **24**, 156–159 (1970).
3. A. Ashkin and J. M. Dziedzic, “Observation of resonances in the radiation pressure on dielectric spheres,” Phys. Rev. Lett. **38**, 1351–1354 (1977).
4. J. E. Bjorkholm, R. R. Freeman, A. Ashkin, and D. B. Pearson, “Observation of focusing of neutral atoms by the dipole forces of resonance-radiation pressure,” Phys. Rev. Lett. **41**, 1361–1364 (1978).
5. K. C. Neuman and S. M. Block, “Optical trapping,” Rev. Sci. Instrum. **75**, 2787–2809 (2004).
6. K. Dholakia, P. Reece, and M. Gu, “Optical micromanipulation,” Chem. Soc. Rev. **37**, 42–55 (2008).
7. D. G. Grier, “A revolution in optical micromanipulation,” Nature **424**, 810–816 (2003).

1975–1990

High-Power, Reliable Diode Lasers and Arrays

Dan Botez

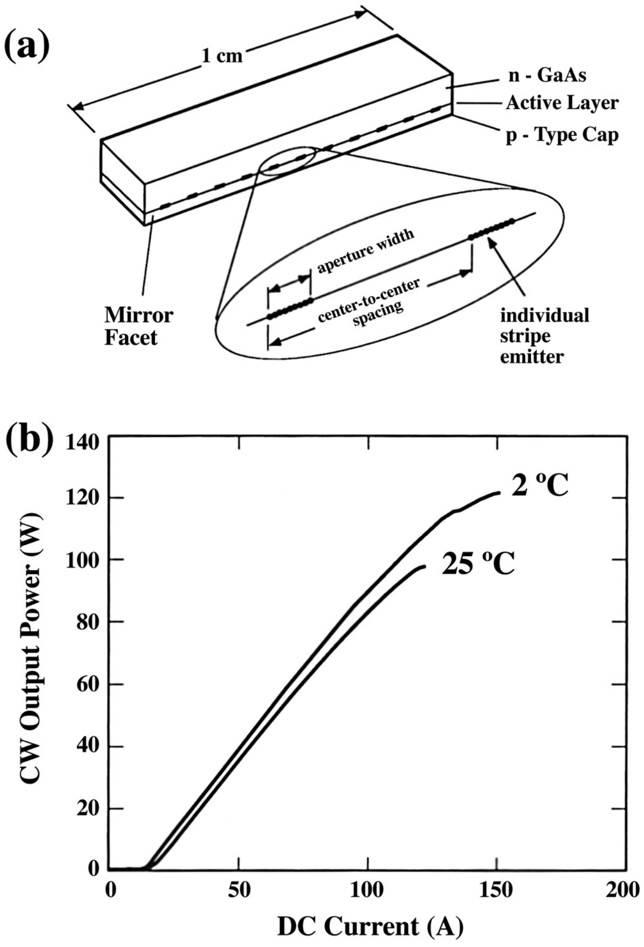
The long-lived diode lasers demonstrated at Bell Laboratories in 1977 produced only a couple of milliwatts (mWs), good enough for fiber-optical communications and later for compact disc reading. Other applications, such as high-speed optical recording, required quasi-continuous-wave (CW) powers in the 50–100-mW range delivered reliably in a single spatial mode.

Since the reliable power is closely related to the optical power density that can damage the emitting facet, designs were needed for enlarging the laser spot size both transversely (i.e., in a direction perpendicular to the plane of the grown layers) and laterally, while maintaining a single spatial mode. In conventional double-heterojunction devices, for which single transverse optical-mode operation is ensured, the main challenge was to create single-mode structures of large lateral spot size. This was realized by introducing mode-dependent radiation losses in so-called antiguided structures, in either the lateral or the transverse directions, on both sides of the defined lateral waveguide. Laterally antiguided diode lasers [1] emitting single-mode peak powers in the 50–80-mW range at 20%–50% duty cycle enabled RCA Laboratories in 1980 to realize high-speed optical recording. At about the same time, Hitachi Central Research Laboratory reported single-mode CW powers as high as 40 mW employing optimized transversely antiguided double-heterojunction devices [2].

In 1980 a breakthrough occurred in high-power diode-laser design with the implementation of the large-optical-cavity concept for increased spot size in the transverse direction [1]. These structures provided transverse spot sizes about 60% larger than in double-heterojunction devices, enabling record-high reliable powers [1]. As a result, the constricted double-heterojunction, large-optical-cavity laser became the most powerful single-mode commercially available diode laser between 1981 and 1986.

In the 1980s the maximum reliable CW power was only about 25% of the maximum achievable power set by catastrophic optical-mirror damage. Mirror damage in diode lasers is caused by thermal runaway at the mirror facets due to increased light absorption and nonradiative recombination with increased drive current [3]. Solutions to suppressing damage required nonabsorbing regions at the mirror facets. As early as 1978 researchers from NEC Laboratories showed that Zn diffusion provides nonabsorbing regions at the mirror facets. This led to a fourfold increase in the maximum achievable CW output power. Then, in 1984 researchers from RCA Laboratories demonstrated mirror-damage suppression by creating, via a single etch-and-regrowth cycle, two-dimensional (2D) waveguiding structures at the mirror facets that were transparent to the laser light. Those devices [3] provided peak output powers of 1.5 W, a fourfold increase over the highest previously reported. However, the early nonabsorbing-mirror approaches were impractical to implement. It took over five years before practical nonabsorbing-mirror lasers were developed and became commercially available.

Around 1982, interest arose in replacing flashlamps with diode-laser arrays as pumps for solid-state lasers. This drive picked up steam with the advent of quantum-well diode lasers since much lower threshold currents could be achieved than in standard double-heterojunction lasers. In early 1983 researchers from Xerox PARC reported very high (>2.5 W CW) CW power quantum-well lasers with optimized facet coatings [3]. Thus, they achieved an eightfold



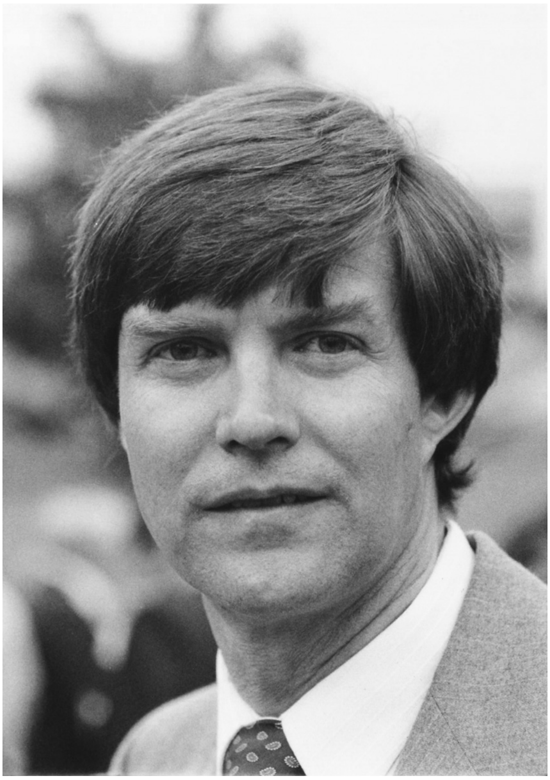
▲ **Fig. 1.** First diode-laser bar to emit 100-W CW power at room temperature: (a) schematic representation, (b) CW output as a function of drive current. (D. R. Scifres and H. H. Kung, "High-power diode laser arrays and their reliability," Chap. 7 in *Diode Laser Arrays*, D. Botez and D. R. Scifres, eds. [Cambridge University Press, 1994].)

increase over the maximum CW power reported from double-heterojunction lasers due both to the use of quantum wells and the use of low-reflectivity dielectric facet coatings. The facet coatings also prevented attack and erosion of the cleaved facets in air, enhancing device reliability [3]. By the mid-1980s large-aperture, high-power, reliable diodes lasting at least 10,000 hours became commercially available [3] from Spectra Diode Laboratories Inc., a start-up company spun off from Xerox PARC. Later that decade, quantum-well laser optimization employing a single, thin quantum well in a large-optical-cavity confinement structure resulted in front-facet, maximum CW wall-plug efficiency as high as 55%.

Quantum-well lasers turned out to offer a solution for practical nonabsorbing-mirror lasers. Researchers from the University of Illinois at Urbana discovered that impurity diffusion causes lattice disordering of multi-quantum-well structures, leading to structures of higher bandgap energy than the energy of light generated in undisturbed multi-quantum-well structures [4]. In 1986, by using impurity-induced disordering, researchers from Xerox PARC and Spectra Diode Laboratories achieved nonabsorbing-mirror structures at the mirror facets [4]. This led to dramatic improvement in maximum CW power from large-aperture devices and was reflected in similar improvements in the reliable CW power output from single-mode devices. This approach led in 1990 to the first 100-mW CW commercially available single-mode diode laser. An alternative nonabsorbing-mirror approach was developed at IBM Zurich Laboratories [3,4]. This approach, called the E2 process, consisting of complete device-facet passivation via *in situ* bar cleaving in ultrahigh vacuum and deposition of a proprietary facet-passivation layer, led to reliable operation of single-mode AlGaAs lasers at 200-mW output power. Today these are the two main nonabsorbing-mirror approaches for multi-watt, reliable operation of both single-stripe lasers and laser bars.

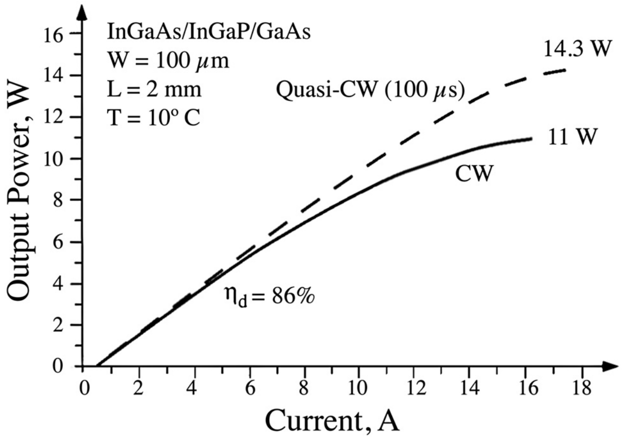
In the early 1990s single-stripe laser and laser-bar development for pumping solid-state lasers started in earnest. For single-stripe, facet-passivated devices of ~400-μm-wide aperture, Spectra Diode Laboratories reported maximum CW power of 11.4 W with reliable CW power of ~4 W [3]. Monolithic laser bars [Fig. 1(a)] composed of an array of 80 separate facet-passivated lasers emitted 100-W CW at room temperature [Fig. 1(b)] [3]. Laser-bar operation in quasi-CW mode at low duty cycles allowed effective heat removal; thus, permitting maximization of the energy per pulse and consequently quite suitable for pumping solid-state lasers. Researchers from Lawrence Livermore National Laboratories (LLNL) reported highly stable high-peak-power, quasi-CW operation after 1 billion shots from 1-cm-long bars [4]. Laser bars were further stacked in 2D arrays to deliver the high powers needed for effective solid-state-laser pumping. Heat removal was a challenging task, and several approaches were developed [3,4]. At the time, the most efficient way to remove heat from 2D arrays was the silicon-based micro-channel cooling technology developed at LLNL [4]. Using that technology LLNL demonstrated 41-bar stacks delivering 3.75-kW peak power [4]. By the end of the decade steady development led to significantly improved performance.

Spectra Diode Laboratories was at the forefront of commercializing high-power diode-laser bars. Donald R. Scifres, the CEO of Spectra Diode Laboratories, was recognized by The OSA in 1996, when he was the awarded the Edwin H. Land Medal for his pioneering scientific and entrepreneurial contributions to the field of high-power semiconductor lasers (Fig. 2). A year later, Dr. Scifres and his wife, Carol, endowed the OSA Nick Holonyak, Jr. Award, dedicated to recognize individuals who have made significant contributions to optics based on semiconductor-based optical devices and materials, including basic science and technological applications.



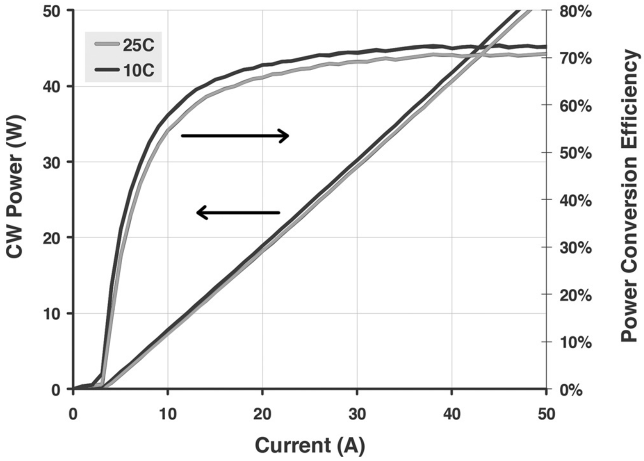
▲ **Fig. 2.** Donald R. Scifres, recipient of the 1996 Edwin Land Medal (at the time). (Courtesy of Dr. W. John Tomlinson, Princeton, New Jersey.)

In the mid-1990s two major developments led to significant increases in the output powers of single-stripe diode lasers: the broad-waveguide-device concept and the use of Al-free, active-region structures. The broad-waveguide concept for asymmetric and symmetric structures involved a large-optical-cavity structure of large equivalent (transverse) spot size as well as low internal cavity loss [5]. The total thickness of the optical-confinement layer of the broad-waveguide structure is quite large, while making sure that lasing of high-order transverse modes is suppressed via losses to the metal contact [5]. Diodes capable of over 10-W CW power were achieved using active regions composed of Al-free, indium (In)-containing material with relatively high mirror-damage power density (see Fig. 3). Later it became clear that adding In to the active-region material significantly decreases the surface recombination velocity, which in turn increases the mirror-damage power density [4]. Indium had another highly beneficial effect with respect to laser-device reliability: it was found to suppress crystal-defect propagation in GaAs-based lasers [4]. That is why currently the most reliable 0.81-μm emitting devices have either InGaAsP or InAlGaAs active regions.



▲ **Fig. 3.** Light-current characteristics in CW and quasi-CW operation for the first single-stripe (100-μm-wide aperture) diode laser to emit over 10-W CW power. (Reproduced with permission from A. Al-Muhanna, L. J. Mawst, D. Botez, D. Z. Garbuzov, R. U. Martinelli, and J. C. Connolly, "High-power (>10 W) continuous-wave operation from 100-μm-aperture 0.97-μm-emitting Al-free diode lasers," Appl. Phys. Lett. **73**(9), 1182 [1998].)

Another key issue that was tackled in the mid-1990s was suppression of carrier leakage out of the lasers' active regions. Since carrier leakage is a thermally activated effect a substantial amount of it causes a significant decrease in the laser slope efficiency as the heat-sink temperature increases. This decrease in slope efficiency is characterized by a



▲ **Fig. 4.** Light-current characteristics and wall-plug efficiency for the first diode-laser bar to emit with over 70% CW wall-plug efficiency at room temperature. (Reproduced by permission of the Institution of Engineering & Technology. Full acknowledgment to M. Kanskar, T. Earles, T. J. Goodnough, E. Stiers, D. Botez, and L. J. Mawst, "73% CW power conversion efficiency at 50 W from 970 nm diode laser bars," Electron. Lett. **41**(5), 245–247 [2005].)

temperature coefficient T_1 [5]. When carrier leakage is suppressed via bandgap engineering, the T_1 parameter has a high value, which reduces the active-region heating [5] and increases the maximum achievable CW power. A high T_1 value also leads to reduced mirror-facet heating; thus, it results in high mirror-damage power-density values [5] and subsequently long-term reliable operation at high CW power levels.

To minimize heating in diode lasers and decrease the heat load as well as improve the lasers' reliability in CW operation the value of the electrical-to-optical power conversion efficiency, the so-called wall-plug efficiency η_p needed to be increased. In 1996, by using broad-waveguide structures with suppressed carrier leakage [6], researchers at the University of Wisconsin–Madison achieved η_p as high as 66%. At the time it was noticed that the devices could not reach their ultimate maximum η_p value due to a built-in voltage differential in the laser structure. Efforts to increase η_p re-started in 2003. By 2005, reductions in the built-in voltage differential as well as laser-structure optimization led to CW wall-plug efficiencies of 73%–75% for laser bars from Alfalight, Inc., nLight Inc., and JDSU Corp. A typical result is shown in Fig. 4, which shows a 50-W CW output delivered with 73% wall-plug efficiency, at 0.97 μm from a 1-cm-wide laser bar [7]. The achievement of record-high wall-plug efficiency was quite a significant development in that it led the typical η_p of commercial laser bars to increase from ~45% to ~65%. Consequently, the dissipated heat that needed to be removed was reduced by more than a factor of 2, which is very important since thermal load management drives the packaged laser weight.

With the advent of the "telecom bubble," feverish activity started around 1999 to create single-spatial-mode, high-power (~1-W CW) 0.98-μm emitting diode lasers for use as pumps for erbium-doped fiber amplifiers to be employed as signal boosters in long-distance fiber-optical communications. Although many complex and elegant approaches were tried, in the end, facet-passivated, 4–5-μm-wide conventional ridge-guide devices prevailed [8]. Even though single-spatial-mode CW powers as high as 1 W are achievable, reliability limits output to ~0.7 W CW due to bulk degradation [8].

Attempts to achieve long-term, reliable operation at higher coherent CW powers by using unstable resonator or master oscillator–power amplifier semiconductor-based configurations have failed [2, 9]. Other approaches consisted of incorporating periodic features, such as distributed-feedback gratings, in the device structure to realize so-called photonic-crystal lasers. However, when using photonic-crystal distributed-feedback devices, the induced periodic refractive-index steps are so small that they are comparable to thermally induced index steps in quasi-CW or CW operation. In turn, these lasers perform well only in low-duty-cycle (≤1%) pulsed operation; thus they are impractical since most applications require high average powers. Only high-index-contrast photonic-crystal lasers that possess long-range coupling between the photonic-crystal sites [2, 9] appear, at present, as the solution to achieving multi-watt CW coherent power from monolithic semiconductor lasers. High-index-contrast, long-range coupling photonic-crystal lasers were realized [9] as early as 1989 in the form of laterally resonant, phase-locked arrays of antiguided lasers, so-called resonant-optical-waveguide arrays. The lateral resonance feature ensures strong coupling between all array elements, in spite of large built-in index steps [9]. In 1991 the resonant-optical-waveguide array became the first diode laser to demonstrate 1-W peak power in a diffraction-limited beam [9], and in 1992 it was theoretically shown to be equivalent to a lateral distributed-feedback structure for which both index and gain vary periodically [9]: that is, an *active* photonic-crystal laser structure. Thus, the resonant-optical-waveguide array did constitute the first photonic-crystal laser developed for high-power, single-mode operation from large-aperture semiconductor lasers. In 1999, resonant-optical-waveguide arrays of an index step more than an order of magnitude larger than in photonic-crystal distributed-feedback structures demonstrated 1.6-W CW power [10] in a nearly diffraction-limited beam from a 200-μm-wide aperture. In 2010, the OSA presented Dan Botez, Philip Dunham Reed Professor at the University of Wisconsin–Madison and co-founder of Alfalight Inc., the Nick Holonyak, Jr. Award for the achievement of active photonic-crystal semiconductor-laser structures for high-coherent-power generation (Fig. 5).

▲ **Fig. 5.** Dan Botez, recipient of the 2010 Nick Holonyak, Jr. Award (at the time). (OPN June 2010 Optical Society Awards.)

High-power, reliable diode-laser technology reached a high degree of maturity by about 2005. Single-stripe devices with 10-W reliable output power and wall-plug efficiencies of ~65% are available from diode-laser manufacturers for various applications including single-diode pumping of solid-state lasers and fiber lasers. Laser bars, used mostly for pumping solid-state lasers, are commercially available with 200-W CW output powers, ~65% wall-plug efficiency and are guaranteed to operate for 30,000 hours. Future developments may involve the commercial realization of active photonic-crystal lasers for watt-range coherent CW powers as well as the use of photonic-crystal structures for emission of the generated light through the substrate (i.e., surface emission) for even higher coherent powers delivered in a reliable fashion.

References

1. D. Botez, D. J. Channin, and M. Ettenberg, "High-power single-mode AlGaAs laser diodes," Opt. Eng. **21**(6), 216066 (1982).
2. N. W. Carlson, *Monolithic Diode-Laser Arrays* (Springer-Verlag, 1994).
3. D. R. Scifres and H. H. Kung, "High-power diode laser arrays and their reliability," Chap. 7 in *Diode Laser Arrays*, D. Botez and D. R. Scifres, eds. (Cambridge University Press, 1994).
4. R. Solarz, R. Beach, B. Bennett, B. Freitas, M. Emanuel, G. Albrecht, B. Comaskey, S. Sutton, and W. Krupke, "High-average-power semiconductor laser arrays and laser array packaging with an emphasis on pumping solid state lasers," Chap. 6 in *Diode Laser Arrays*, D. Botez and D. R. Scifres, eds. (Cambridge University Press, 1994).
5. D. Botez, "Design considerations and analytical approximations for high continuous-wave power, broad-waveguide diode lasers," Appl. Phys. Lett. **74**(21), 3102–3104 (1999).
6. D. Botez, "High-power Al-free coherent and incoherent diode lasers," Proc. SPIE **3628**, 2–10 (1999).
7. M. Kanskar, T. Earles, T. J. Goodnough, E. Stiers, D. Botez, and L. J. Mawst, "73% CW power conversion efficiency at 50 W from 970 nm diode laser bars," Electron. Lett. **41**(5), 245–247 (2005).
8. G. Yang, G. M. Smith, M. K. Davis, D. A. S. Loeber, M. Hu, Chung-en Zah, and R. Bhat, "Highly reliable high-power 980-nm pump laser," IEEE Photon. Tech. Lett. **16**(11), 2403–2405 (2004).
9. D. Botez, "Monolithic phase-locked semiconductor laser arrays," Chap. 1 in *Diode Laser Arrays*, D. Botez and D. R. Scifres, eds. (Cambridge University Press, 1994).
10. H. Yang, L. J. Mawst, and D. Botez, "1.6 W continuous-wave coherent power from large-index-step ($\Delta n \approx 0.1$) near-resonant, antiguided diode laser arrays," Appl. Phys. Lett. **76**(10), 1219–1221 (2000).

Tunable Solid State Lasers

Peter F. Moulton

While the wavelength of any laser can be varied, lasers get classified as tunable when their tuning range becomes a substantial fraction of their center wavelength. Despite having lower optical gain than narrow-line rare-earth doped crystal lasers such as Nd^{3+}-doped YAG, tunable lasers are desirable for a number of reasons. In laser-based spectroscopy, laser tuning allows one to access spectral features of interest, while in laser propagation through the atmosphere, tuning can be used to avoid atmospheric absorption lines. A large tuning range implies the ability to generate and amplify short pulses of light. The development of practical and efficient tunable solid state lasers has led to a scientific revolution and an emerging industrial revolution in laser processing of materials, based on the generation of electromagnetic pulses with femtosecond and recently attosecond duration.

Most broadly tunable lasers employ ions from the "3d" portion of the periodic table. Figure 1 presents so-called configuration-coordinate diagrams that help explain the broad tunability of 3d ions. The diagrams are a greatly simplified schematic representation of the combined energy of the laser-active ion and its environment as a function of the positions of the atoms surrounding the ion. In equilibrium, the overall energy is minimized, and the system energy increases as the coordinate deviates from the equilibrium position. Deviation occurs as a result of the always present vibrations of the atoms, which appear even at the lowest temperatures from the uncertainty principle of quantum mechanics. The left-hand diagram shows the case where, when the ion energy level changes from a "ground" state to an "excited" state, the equilibrium position for the configuration coordinate is unchanged. The right-hand side shows the case where the equilibrium position does change.

An important concept regarding the linewidth of the transitions between the ion ground and excited states is the Franck–Condon principle. Stated in classical terms, when an active ion undergoes a transition, it occurs so quickly that the atomic surroundings do not move, as shown by the vertical arrows in the diagrams. The left-hand diagram is representative of the type of narrow-linewidth transitions among levels of the rare-earth ions, since changing the electronic state of the spatially compact wavefunctions of the rare earths has negligible effect on the surrounding atoms.

The electronic wavefunctions of 3d ions have a larger spatial extent than those of rare-earths and have a stronger interaction with their environment. The case illustrated in the right-hand diagram shows what happens with a strong interaction, exciting the electronic level leads to a new equilibrium position. As is evident from the arrows, the energy associated with ground-to-excited-state transitions does vary with the displacement, leading to a large spread in energies and hence a large linewidth. The energies for the absorption (ground-to-excited) transitions are generally distinct and higher than those for emission and possible laser operation (excited-to-ground transitions). As a result, even with only two electronic transitions, one can observe four-level laser operation (as shown by the numbers in Fig. 1) as the peak absorption and emission wavelengths do not overlap. These types of transitions are often referred to as "vibronic," a concatenation of vibrational and electronic.

After the demonstration of the ruby laser, and around the same time as the development of rare-earth-doped lasers, there were demonstrations of the first broadly tunable solid state lasers, based on 3d-ion transitions. In particular, in 1963 L. F. Johnson and co-workers at Bell Labs

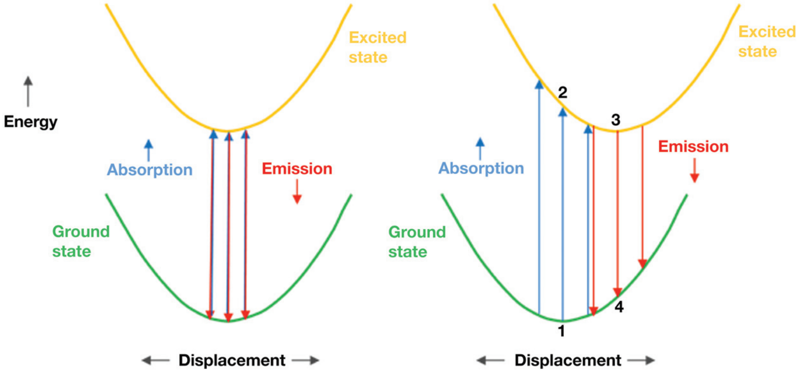
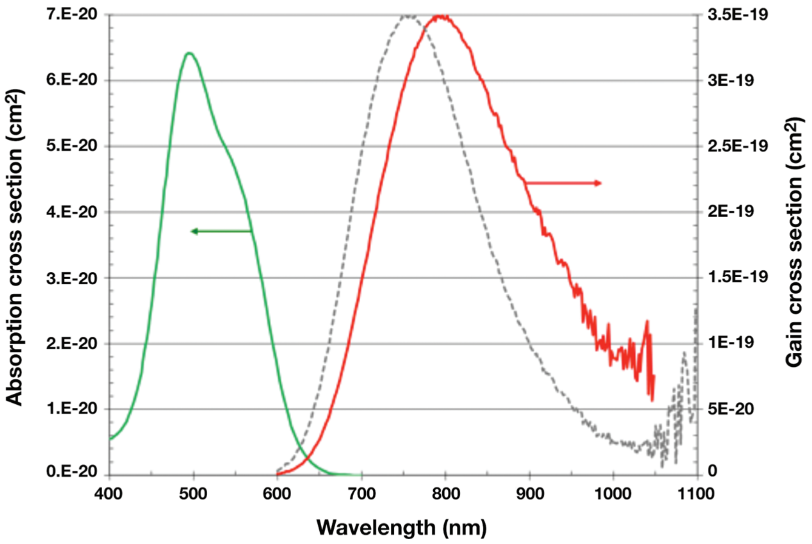


Fig. 1. Configuration-coordinate diagrams for two cases of paramagnetic-ion transitions.

reported "optical maser oscillation from Ni^{2+} in MgF_2 involving simultaneous emission of phonons," which, translated to now-accepted terminology, would be "laser operation on vibronic transitions." Subsequent work by the same group showed operation on vibronic transitions from Co^{2+} ion in MgF_2 and ZnF_2 around 1750–2150 nm, prism-based tuning, albeit in discontinuous segments from Ni:MgF_2, and operation from V^{2+}-doped MgF_2 around 1100 nm. The major drawback to these first vibronic lasers was that, because of thermally induced non-radiative processes, relatively low-threshold operation with lamp pumping required cooling of the laser crystals to cryogenic temperatures. The author, working at MIT Lincoln Laboratory in the 1970s, became aware of the early Bell Labs work and realized that the use of lasers, rather than lamps, as pump sources could greatly reduce the engineering complexity of the systems. In particular, Nd-doped solid state lasers operating around 1300 nm proved effective in pumping both Ni^{2+} and Co:MgF_2 lasers. In the subsequent work, he had some success with the Co:MgF_2 laser, which proved capable of tuning from 1630–2080 nm at LN_2 temperatures and 1750–2500 nm at room temperature. Other 3d systems he studied showed clear evidence of a problem that has plagued many tunable solid state lasers: excited-state absorption (ESA). For most ions there are a number of 3d levels above the first excited state, i.e., the upper laser level. Depending on the positions of the levels in the configuration-coordinate diagram, it is possible that, for the desired laser wavelength, induced transitions to one or several of these levels may be possible. The net cross section that determines the laser gain is the cross section for transitions to the lower laser level minus the cross section for transitions to the higher-lying states, and this reduces laser efficiency and can even prevent laser operation.

The announcement of room-temperature, 750-nm-wavelength-region, tunable laser operation from Cr^{3+}-doped $BeAl_2O_4$ (alexandrite) in 1979 re-ignited interest in Cr^{3+}-doped lasers beyond ruby. At first, laser operation was thought to be, like ruby, on a narrow-line transition but spectroscopic investigation showed that it was in fact a vibronic. However, the gain in alexandrite lasers is relatively low, limiting applications, and today the most widespread use of alexandrite is in lamp-pumped, long-pulse lasers used for a variety of medical applications. The majority of other Cr^{3+}-doped tunable materials studied showed low conversion of pump to laser power, generally attributed to ESA. One class (colquirite structure) of materials, first developed at Lawrence Livermore National Laboratory, includes the crystals $LiCaAlF_6$ (LiCAF) and $LiSrAlF_6$ (LiSAF) and was shown to have relatively weak ESA and thus high efficiency. However, the thermo-mechanical properties of the colquirite host crystals (with thermal conductivities 10%–20% of the sapphire and alexandrite host crystals) significantly limit their ability to generate high average powers free of significant thermo-optic distortion of the output beam and, ultimately, free of fracture to the laser material.

While listening to a presentation on a particular type of color-center laser the author noted the simplicity of that system: there were no excited states above the upper laser level that could cause ESA. A subsequent review of the periodic table showed that one 3d ion, Ti^{3+}, has only a single 3d electron. The five-fold degenerate free-space state for that electron placed in a typical crystal, to first order, splits into a three-fold degenerate ground state, $^2T_2$, and a doubly degenerate upper state, ^2E. Any higher-lying states result from transitions that take the single electron out of the 3d shell and



◀ **Fig. 2.** Absorption (green) and emission (red) cross sections for Ti:sapphire and a relative plot (dashed gray curve) of the measured fluorescence spectrum. The noise in the long-wavelength region is from the detection system.

could be so high in energy as to not create ESA. There were reports on the basic spectroscopy of Ti^{3+}-doped Al_2O_3 (Ti:sapphire) with data on absorption and fluorescence. Given the superior thermomechanical properties of sapphire, proven with the ruby laser, it looked to be a good choice for a Ti^{3+} host.

The author obtained crystal samples from Robert Coble's group at MIT, where they had been studying the diffusion of oxygen in sapphire by using the oxidation state of Ti as a tracer. (Coble was the developer of the first transparent ceramics, paving the way for sodium arc lamps and, later, laser-quality ceramics.) The author's measurements of the absorption cross section and fluorescence spectra, shown in Fig. 2, showed much broader emission than earlier reports. When one converts the emission data to gain cross section (also plotted in Fig. 2), multiplying by the necessary $(\text{wavelength})^5$ correction, the tuning range is unusually broad. The spectral breadth of the emission results, in part, from Jahn–Teller splitting of both the ground and upper levels of the ion, leading to a more complicated configuration coordinate than shown in Fig. 1, where both the ground and excited states have multiple-energy versus displacement curves. The author also determined the fluorescence decay time and found a room-temperature value of 3.15 μs. The short lifetime seemed to indicate low quantum efficiency, but if one estimates the radiative value based on the strength of the measured absorption in the material, as well as on optical gain measurements, one finds high quantum efficiency, on the order of 80% at room temperature. The short lifetime and associated high gain cross section (in the range $3\text{–}4 \times 10^{-19}$ cm^2) result from the trigonal symmetry of the Ti^{3+} site in sapphire, which acts to strongly activate the dipole-forbidden $^2E \rightarrow {}^2T_2$ transitions.

The author first obtained laser operation from the material in May of 1982 and reported the results in June at the Twelfth International Quantum Electronics Conference in Munich. There was a delay in publication in a fully refereed journal until 1985 while the author worked, unsuccessfully, to patent the system, became engaged in other technical work, and left MIT Lincoln Laboratory to help start a company. The results published in 1985 included demonstrations of pulsed laser operation with lamp-pumped, dye-laser pumps, frequency-doubled, Q-switched, Nd:YAG laser pumps, and continuous-wave (CW) operation with argon-ion-laser pumps, with cryogenic cooling used to obtain true-CW operation. In laser operation, tuning experiments showed that the observed tuning range and that predicted by fluorescence measurements were in good agreement, confirming that ESA was not a factor in laser operation.

The first commercial Ti:sapphire laser product, an argon-ion-laser pumped CW device, was introduced by Spectra-Physics in 1988 and was followed shortly after by one from the author's company, Schwartz Electro-Optics, that included an option for a single-frequency, ring-laser configuration. Early applications of the products included use as a diode-laser substitute in the development of other solid state lasers, notably Er-doped fiber amplifiers pumped at 980 nm for telecom applications and later Yb-doped, high-efficiency crystal lasers. With the discovery that nonlinear effects in the CW Ti:sapphire laser crystal, namely Kerr-effect lensing, could lead to generation of 60-fs-duration pulses, the utility of Ti:sapphire lasers greatly expanded. The irony of this is that the nonlinearity in the solid state laser medium might have been expected to be a limit to the mode-locking properties of the system, but it in fact provided a path to generation of femtosecond pulses. Subsequent technology improvements, including dispersion-compensating intracavity elements, broadband mirrors with appropriate optical dispersion and phase characteristics, and sophisticated pulse diagnostics, led to direct generation of 3.6-fs-duration pulses at 800 nm, slightly more than one optical cycle. These are claimed to be the shortest pulses directly generated by any laser system and close to the limit expected from the 100-THz gain bandwidth of Ti:sapphire. Commercial mode-locked Ti:sapphire lasers emerged in 1991 with picosecond-duration pulses, followed shortly by Kerr-lens-based systems providing 100-fs-duration pulses, and brought reliable ultrafast-laser technology to a broader base of users, replacing dye-laser-based sources that required long setup times with "turn-key" sources that allowed users to devote more time to science and much less to laser maintenance. The high Ti:sapphire laser gain cross section yields a pulse saturation fluence on the order of 0.8 J/cm^2, comparable with 0.7 J/cm^2 of Nd:YAG-generated high-energy pulses. If one uses a Q-switched, frequency-doubled, Nd:YAG solid state pump laser, the Ti:sapphire medium will be able to integrate and store the pump energy, which must then be extracted within a microsecond or so of the pump pulse.

The combination of femtosecond-duration pulses produced by CW Ti:sapphire lasers and high-gain, high-energy amplifiers pumped by pulsed, Q-switched lasers has led to widely used systems for high-intensity pulse generation. A key technology for this combination is the chirped-pulse amplification (CPA) technique of Strickland and Mourou, first reported in 1985 and nicely matched to the properties of Ti:sapphire. With the availability of large-aperture Ti:sapphire crystals, the ultimate limit on energy is set by the pump laser, and the limit on pulse rate is set by a combination of the pump laser and thermal effects in the Ti:sapphire material. At present, regenerative systems are widely available on a commercial basis, with pulse energies of tens of millijoules and pulsewidths <40 fs, with cryogenic cooling used for systems producing 20–30 W average power. In sum, commercial sales of Ti:sapphire lasers to date, including associated green-wavelength pump lasers, are on the order of $1 billion, not counting very-high-power systems installed or being built at major research laboratories.

At this writing, there is active development to scale up the peak power/energy of Ti:sapphire CPA systems. The highest reported power is 2×10^{15} W (2 PW), from a system in Shanghai, with a final stage pumped by a Nd:glass laser providing 140 J of 527 nm pump energy. The APPOLON Ti:sapphire laser system, under construction in France, has a goal of 10 PW in a pulse of 150 J in 15 ps. Figure 3 shows the pumped final stage of the one the Gemini amplifiers at the Central Laser Facility (CLF), Rutherford Appleton Laboratory, Oxford, UK, generating 25 J of pulse energy in a 30-fs pulse.

Key new advances in tunable solid state lasers are now almost entirely driven by their application to ultrafast pulse generation and include diode-pumped, rare-earth, Yb-doped crystals that can generate pulses on the order of 100 fs, and Cr^{2+}-doped ZnSe and similar II-VI

▲ **Fig. 3.** One of the two Gemini amplifiers at Rutherford Appleton Laboratory, showing, in the center, a green-laser-pumped Ti:sapphire crystal 90 mm in diameter and 25 mm thick. With 60 J of pump energy the system has generated 25 J of output energy in a 30 fs pulse. (STFC Gemini Laser Facility/Chris Hooker).

semiconductor hosts, providing high-gain operation similar to that of Ti:sapphire lasers but centered at 2500 nm. The longer wavelength is of great interest for attosecond pulse generation in the x-ray wavelength region through high-harmonic generation. The limited number of pages for this article requires that we leave out further discussion of these developments in tunable solid state lasers.

Other articles in this book provide details of the exciting science and Nobel-prize-winning work that has been enabled by the development of tunable solid state lasers.

1975–1990

Ultrashort-Pulse Lasers

Erich P. Ippen

Introduction

A particularly remarkable aspect of lasers is their ability to emit shorter flashes (pulses) of light than achievable with any other means. This ability has, over the years, advanced the observation and measurement of events from the nanosecond timescale down to the picosecond (10^{-12}), femtosecond (10^{-15}), and even attosecond (10^{-18}) timescales. To use such pulses has required the development of new methods for measuring and characterizing the pulses themselves on ultrafast timescales beyond the reach of electronics. These methods have, in turn, made it possible to study ultrafast phenomena in ways that produced completely new insights into the evolution of such phenomena in physics, chemistry, and biology [1]. As ultrashort-pulse laser technology has developed, its other characteristics such as the high peak power and ultrabroad bandwidth packed into a short pulse have also found important applications. The compression of even very modest amounts of pulse energy into femtosecond durations produces sufficiently high peak power for precision machining and micro-surgery without unwanted damage and for nondestructive nonlinear methods of microscopy that produce three-dimensional (3D) biological imaging with micrometer resolution. The ultrabroad bandwidths associated with femtosecond pulses have made possible 3D medical imaging via optical coherence tomography (OCT), simultaneous creation of many wavelength-multiplexed optical communication channels with only one source, and major advances in precision spectroscopy and optical clocks [2,3].

The Optical Society (OSA) played a major role in supporting the field, starting with its creation of the first International Conference on Picosecond Phenomena in 1978 (name changed in 1984 to Ultrafast Phenomena to reflect the emergence of femtosecond science and technology). Held every two years since then (for the 19th time in 2014, the year of this writing) with continuing OSA support, this successful conference has provided perhaps the greatest testament to the continuous technological development and widespread impact of the field with its 19-volume series of hardcover proceedings [4]. OSA journals became the primary source of publications on ultrafast optics and photonics. Multiple sessions on ultrafast optics and its applications every year at conferences like CLEO, QELS, IQEC, and OFC have been essential to advancing the technology and its applications to science and engineering.

Flashlamp-Pumped Picosecond Systems

Nd:glass Lasers

The era of ultrashort pulses began in earnest with the demonstrations in the mid-1960s, by DeMaria and co-workers at United Aircraft, of passive (self) mode-locking in a Nd:glass laser. Mode locking was achieved with a cell of absorbing dye inside the laser that was designed to bleach (saturate) sufficiently and rapidly enough to favor transmission of high intensity peaks over continuous emission and, therefore, the development of short pulses. The passive, saturable absorber technique, in various forms, remains the basis for ultrashort-pulse generation today. The mode-locked Nd:glass laser pulses, too short to measure at first, were later verified to be on the order of 5–10 picoseconds in duration. For almost a decade, this laser system dominated and drove the development of ultrashort-pulse technology and its applications. For the second decade and a half, mode-locked dye lasers reigned and pushed pulse durations into the femtosecond domain. Finally, with the emergence of new techniques in the late 1980s, passive mode-locking of solid state lasers regained importance and led to the wide range of compact, robust, femtosecond laser systems we have today.

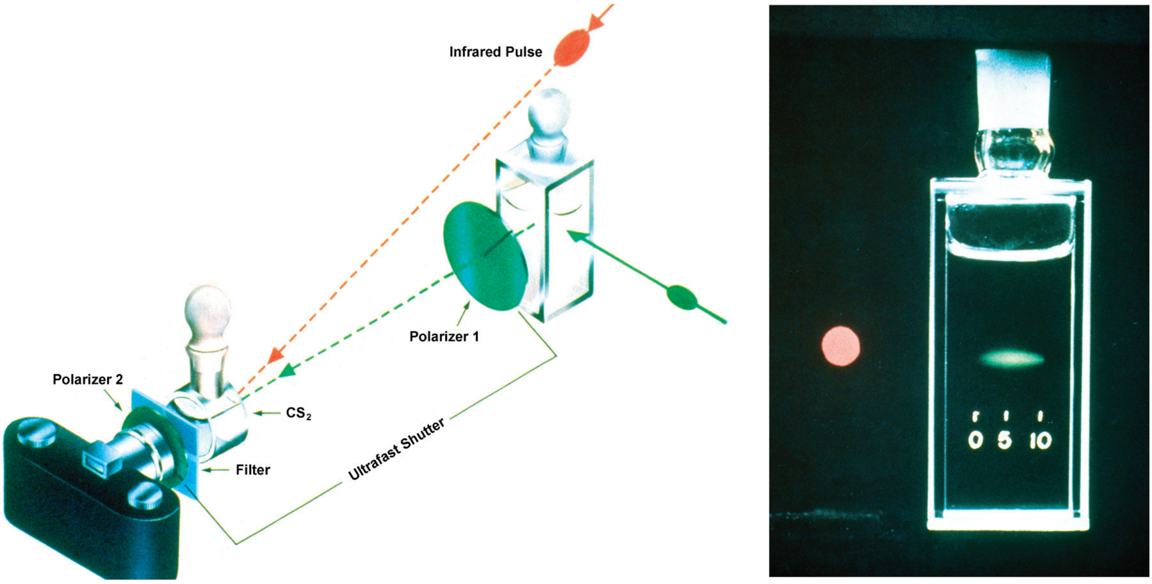
Ultrafast Measurement Techniques and Applications

Stimulated by mode-locked Nd:glass laser demonstrations, many of the ultrashort-pulse characterization, manipulation, and application methods still in use today were invented and developed in the 1960s [5]. Within a year of the invention of the passively mode-locked Nd:glass laser, several methods for pulse measurement with sub-picosecond resolution had been proposed and demonstrated. These techniques essentially use optical pulses to measure themselves. The laser output beam is split into two, one is delayed with respect to the other, and they are combined in a nonlinear crystal to generate second harmonic light (SHG). SHG is a maximum when the two pulses exactly overlap and decreases with delay in either direction. A plot of SHG versus delay yields the second-order autocorrelation function of the pulse intensity I(t). Fitting the observed intensity autocorrelation function to that expected for the pulses requires some assumptions about pulse shape, as this simple method is inherently insensitive to pulse asymmetry. Nevertheless, information about substructure and frequency chirp within the pulse can be deduced by comparing the assumed fit with that expected from the optical frequency spectrum. Methods for complete pulse characterization via frequency-resolved optical gating (FROG) were not developed until the early 1990s. The relatively slow repetition rate of flashlamp-pumped systems made the requirement of repetitive measurements at variable delay somewhat tedious at first. More rapid progress was permitted by the invention of a single-shot method in which two identical copies of a pulse are passed through a two-photon absorbing medium in counterpropagating fashion. The two-photon-induced fluorescence (TPF) intensity pattern, viewed from the side, provides another direct measure of the second-order autocorrelation function. Although widely used and valuable in early work, the TPF method subsequently gave way again to SHG-based methods with the advent of high repetition-rate continuous wave (CW) systems in the mid-1970s.

Most other present-day methods for manipulating pulses and applying them also developed rapidly during this period. It was shown that pairs of gratings can compensate for the chirp produced by linear dispersion in a laser. It followed that pulses could be shortened further by external self-phase modulation followed by a grating pair. Ultrafast responses in materials were observed by splitting a pulse beam into two, an excitation (pump) and a probe, and varying the time delay between them. Continuum generation, discovered by Alfano and Shapiro, made possible the simultaneous probing of changes over broad spectra. The ultrafast optical Kerr shutter, invented by Duguay and co-workers, was used as a picosecond camera to captured dramatic images of light pulses in flight (see Fig. 1 [6,7]) and to carry out the first demonstrations of 3D imaging via variable delay optical gating that later inspired the development by Jim Fujimoto of OCT for medical imaging (see Fig. 2). Other still-useful techniques such as up-conversion gating and transient grating spectroscopy were also demonstrated during this era. Scientific applications expanded to wide-ranging studies of nonlinear optics, picosecond interactions in liquids, and ultrafast processes in chemistry and biology [5].

Pulsed Dye Lasers

Known to have even more broadband potential than the Nd:glass laser, dye lasers were pursued shortly thereafter. The first experiments utilized picosecond pulses from frequency-doubled Nd:glass lasers to generate similarly short pulses from dye lasers. Passive mode-locking of the flashlamp-pumped Rhodamine 6G laser with a saturable dye soon followed. Within a few years the wavelength coverage of ultrashort-pulse dye lasers ranged from almost 400 nm to 1150 nm and amplified peak powers in the



(a) (b)

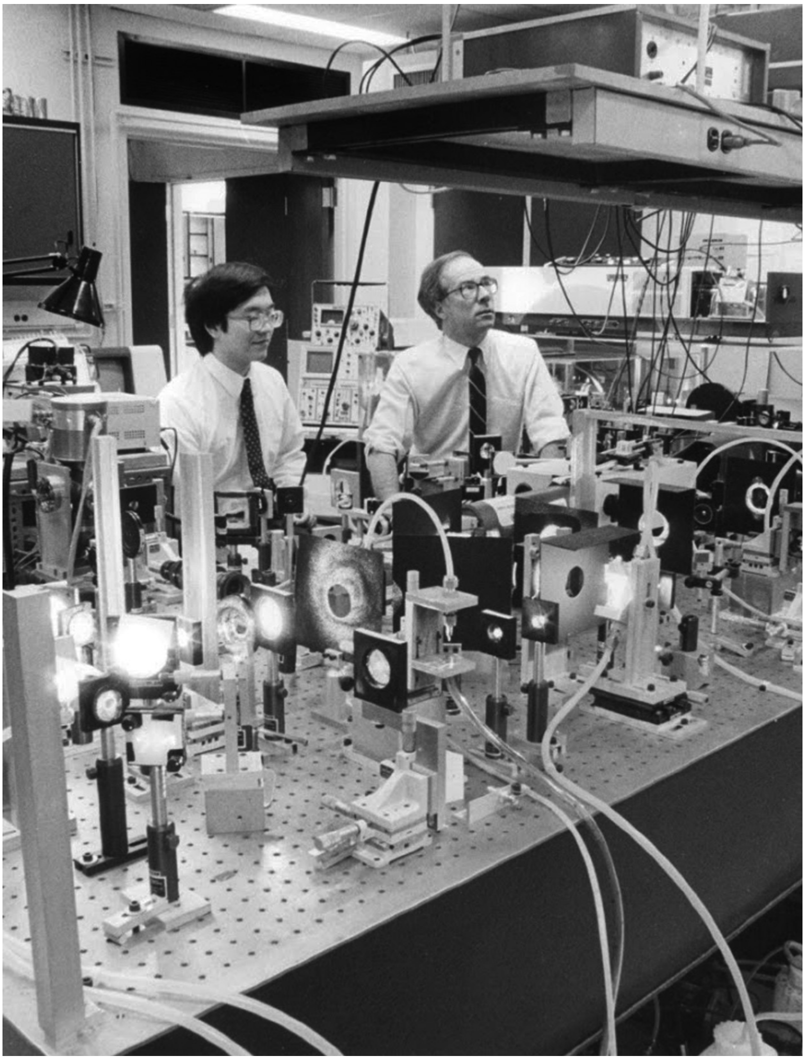
▲ **Fig. 1.** Light in flight. An optical Kerr effect shutter, operated by a picosecond infrared pulse, is used to capture the image of a picosecond pulse passing through a lightly scattering liquid. (a) experimental arrangement (b) the photo. Reprinted with permission from M. A. Duguay and J. W. Hansen, Appl. Phys. Lett. **15**, 192–194 ©1969, AIP Publishing LLC.

gigawatt range had been demonstrated, to a great extent by the Bradley group at Imperial College. As the pulse-forming dynamics of dye systems began to be studied in detail, the following question arose: How were such short pulses generated with saturable absorber dyes having much longer recovery times? In Nd:glass lasers, pulses were shown in studies to build up from noise, with the saturable absorber selecting the most intense pulse and determining the final duration by its recovery time. Dye-laser pulses were getting much shorter. This could happen, according to the insight of G. H. C. New, because, although bleaching the saturable absorber could only shape the leading edge of the pulse shorter than its recovery time, the trailing edge could be shaped by rapid saturation (depletion) of the dye gain medium. By 1975 all of these analyses were put into the subsequently very influential steady-state analytical descriptions, by Haus, of "fast" and "slow" saturable-absorber mode-locking that predicted shapes, durations, and stability [8,9,10].

Continuous-Wave Femtosecond Systems

CW Dye Lasers

Mode-locking of CW dye lasers offered a range of new possibilities for ultrashort-pulse generation. The continuous sources of high-repetition-rate pulses greatly facilitated measurement and the optimization of pulse characteristics via cavity alignment and saturable absorber concentration. With the first reports, in 1972, of passive mode-locking of a CW dye laser, pulses as short as 1.5 ps were reported. Within a year, the first pulses shorter than a picosecond had been produced by Shank and Ippen at Bell Labs (Fig. 3). The femtosecond era had begun. Pulses of 300 fs duration were soon achieved, and application of this new femtosecond capability to studies of ultrafast dynamics in physics, chemistry, and biology followed rapidly. Novel up-conversion pump-probe methods were developed, pulses of 500 fs in duration were amplified to peak powers of gigawatt intensities, and synchronized continuum generation made possible sub-picosecond time-resolved spectroscopy with greatly improved sensitivity and signal-to-noise ratio. Invention of the colliding-pulse

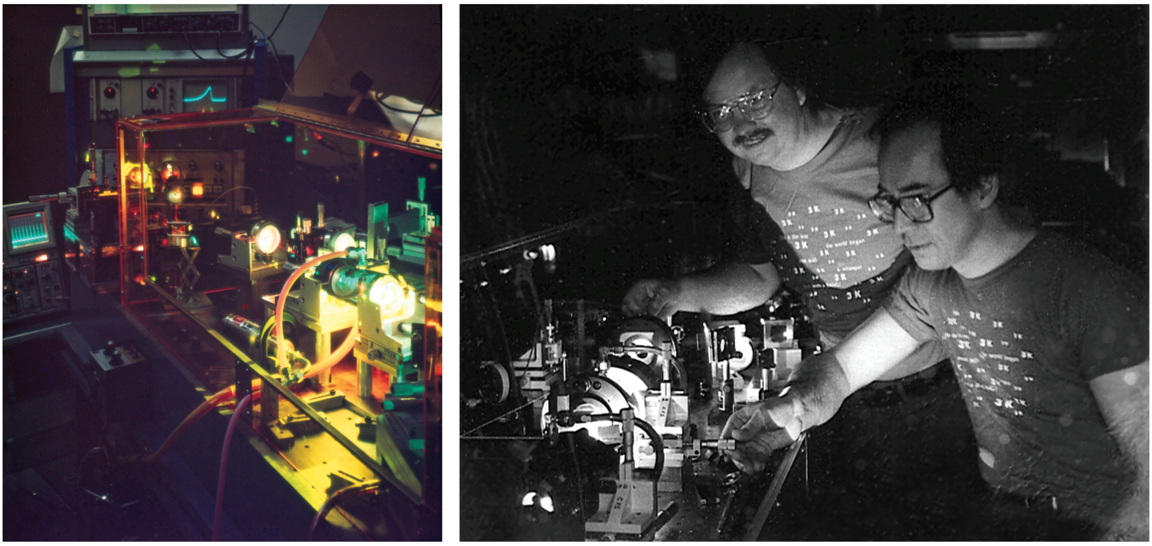


◀ **Fig. 2.** MIT Ultrafast Optics Lab 1985. Erich Ippen and student James Fujimoto view experiment achieving the first demonstration of optical ranging through skin, prelude to the development of Optical Coherence Tomography by James Fujimoto.

mode-locked (CPM) geometry in 1981 at Bell Labs reduced pulse durations to the 100-fs level and further improved stability. The interplay between self-phase modulation and internal dispersion was analyzed theoretically and optimized experimentally via prism pairs to reduce durations further to below 30 fs. Rapid progress was made by several groups, and with amplification and external compression, a record duration of 6 fs, a record that lasted more than a decade, was achieved. Amplified systems, pumped by either 10-Hz frequency-doubled Nd:YAG lasers (Fig. 4) or by kHz copper-vapor lasers, further extended the capability of femtosecond technology and its range of applications. The experiments leading to the 1999 Nobel Prize for chemistry [1] were achieved with this early femtosecond dye-laser technology.

Semiconductor Diode Lasers

Recognized as having gain response times very similar to those of dye lasers, semiconductor diode lasers also became the subject of mode-locking attempts. Shortly after active mode-locking was first demonstrated at MIT in 1978, passive mode-locking of a GaAlAs diode laser in an external cavity produced 5-ps pulse durations at a repetition rate of 850 MHz. Sub-picosecond pulses were later achieved at higher repetition rates, and integrated CPM geometry devices produced pulses as short as 640 fs at a repetition rate of 350 GHz. Impressive demonstrations of high-power, sub-picosecond pulses



(a) (b)

▲ **Fig. 3.** The first femtosecond laser, a Rhodamine 6G dye laser passively mode-locked by a DODCI saturable absorber dye. (a) Instruments record the pulse train and a sub-picosecond-resolution pump-probe trace of a molecular response. (b) Chuck Shank and Erich Ippen with their laser.

were achieved by Delfyett and co-workers with pulse compression and semiconductor optical amplification. Stable, transform-limited pulse generation with semiconductor diodes has, however, for the most part depended on external-cavity-controlled picosecond sources. Pump-probe investigations revealed that ultrafast nonequilibrium carrier dynamics in a semiconductor make the generation of pulses shorter than 1 ps problematic.

Color-Center Lasers

An important capability for early 1.5-μm-wavelength ultrafast research was provided by the CW color-center laser. First mode-locked by synchronous pumping, the KCl color center laser was thrust into further prominence by Mollenauer's demonstration at Bell Labs that it could produce femtosecond pulses by operating as a "soliton laser." This was achieved by coupling the laser output into an anomalously dispersive, soliton-shaping, optical fiber, the output of which was then coupled back into the laser. It was soon discovered, however, that soliton formation in the fiber was not necessary since this coupled-cavity approach also worked with normal-dispersion fiber. Experiments at MIT further revealed the underlying pulse-shortening mechanism to be the interference of each pulse with a copy of itself that had been self-phase modulated in the fiber. This method, dubbed additive-pulse mode locking (APM), was shown to be compatible with the Haus fast-absorber model. Recognized as a means of creating an "artificial" fast absorber out of reactive nonlinearity in a lossless dielectric, APM then stimulated the application of this technique to a variety of other lasers [11].

Fiber Lasers

Interest in fiber lasers developed rapidly after demonstrations at Southhampton of efficient optical amplification in low-loss fibers doped with rare earths. The key mechanism for ultrashort-pulse generation in fiber lasers—nonlinear polarization rotation—was also found to be describable by the fast-absorber model of Haus developed in the context of APM analysis. Earliest progress was made using Nd:fiber lasers, in both actively mode-locked and passively mode-locked configurations. By 1992 pulse durations as short as 38 fs had been generated at 1.06 μm in a Nd:fiber laser utilizing nonlinear polarization rotation and prism pairs for dispersion compensation. By the turn of the century, however,

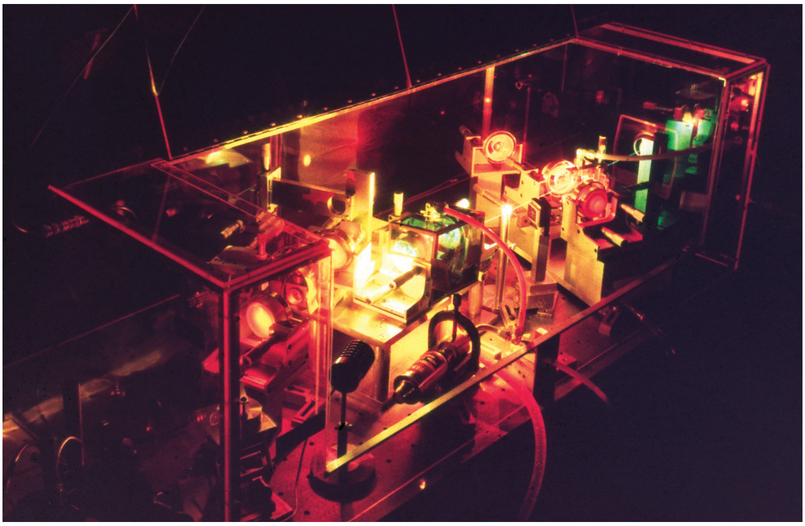


Fig. 4. High power, 3-stage, femtosecond dye laser amplifier pumped by frequency-doubled Nd:YAG laser at 10 Hz.

development of the much more efficient Yb:fiber laser led to considerably higher powers at 1 μm wavelengths, with similarly short pulses and more compact geometries. In the late 1980s the attention of researchers also turned to Er:fiber lasers for wavelengths being used for optical fiber communications and where fibers were anomalously dispersive, permitting soliton pulse shaping and shortening. Sub-picosecond pulses were first achieved, at NRL and at Southhampton in figure-eight geometries that used a nonlinear loop mirror for intensity modulation and pulse stabilization, and then, at MIT and Southhampton, in the ring geometry stabilized by nonlinear polarization rotation that achieved common usage. The MIT stretched-pulse laser achieved shorter pulses and higher pulse energies and was soon commercialized. Although not geared to the high-power applications of Yb:fiber lasers, Er:fiber lasers continue to be pursued for silicon photonics, fiber-based communications, and a variety of eye-safe applications.

Free-space Solid-state Lasers

The discovery of APM and the prospect it offered for CW mode-locked solid-state lasers led to its application to Nd:YAG, Nd:YLF, and Ti:sapphire systems. To permit amplification to high power, Strickland and Mourou in 1985 demonstrated the chirped-pulse amplification (CPA) scheme that would ultimately open the door to attosecond and petawatt optical physics. With the discovery of the Kerr-lens mode-locked (KLM) Ti:sapphire laser in 1991 by the Sibbett group in St. Andrews, KLM became the dominant ultrashort-pulse generation mechanism in free-space solid-state lasers. Femtosecond science and technology entered a new era, one with a wider variety of femtosecond-laser media, shorter pulses, extreme powers, ultrabroad bandwidths, and, quite dramatically, the convergence of ultrashort-pulse lasers with ultranarrow-linewidth lasers, precision spectroscopy, and optical clocks. This modern era is the subject of a following article.

References

1. A. H. Zewail, Nobel Prize in Chemistry, 1999.
2. T. W. Hänsch, Nobel Prize in Physics, 2005.
3. J. L. Hall, Nobel Prize in Physics, 2005.
4. *Ultrafast Phenomena I–XVIII*, Springer Series in Chemical Physics (Springer, 1978–2012).
5. S. L. Shapiro, ed., *Ultrashort Light Pulses*, 2nd ed., Vol. **18** of Topics in Applied Physics, (Springer-Verlag, 1984).
6. M. A. Duguay and J. W. Hansen, "An ultrafast light gate," Appl. Phys. Lett. **15**, 192–194 (1969).

7. M. A. Duguay, “Light photographed in flight: Ultrahigh-speed photographic techniques now give us a portrait of light in flight as it passes through a scattering medium,” Am. Sci. **5**, 550–556 (1971).
8. H. A. Haus, “Theory of mode locking with a fast saturable absorber,” J. Appl. Phys **46**, 3049–3058 (1975).
9. H. A. Haus, “Theory of mode locking with a slow saturable absorber,” IEEE J. Quantum Electron. **11**, 736–746 (1975).
10. E. P. Ippen, “Principles of passive mode locking,” Appl. Phys. B **58**, 159–170 (1994).
11. A. M. Weiner, *Ultrafast Optics* (Wiley, 2009).

1975–1990

Ground-Based Telescopes and Instruments

James Breckinridge

By 1916, the American astronomer George Ellery Hale (see Fig. 1), a founding member of The Optical Society (OSA), had designed and built an optical solar telescope on Mt. Wilson and measured the strength of magnetic fields on the Sun using his new invention: the solar magnetograph. This opened a new era in astronomy and demonstrated to all the merits of adding optical physics to astronomy. In 1916 the Mt. Wilson Observatory, under the direction of Hale, had just completed the 60-inch reflecting telescope and it was becoming productive. Hale hired George Ritchey to figure the 60-inch mirror with a hyperbolic primary and secondary to extend the field of view (FOV) of the standard Cassegrain telescope. Most astronomical telescopes today use this optical configuration.

Hale's career started out at the University of Chicago, where he met A. A. Michelson (OSA Honorary Member) in 1889 when he arrived at the University of Chicago. Hale nominated Michelson for the Nobel Prize in Physics in 1907. In 1916 Hale, director of Mt. Wilson Observatory, was elected vice-president of OSA. Later (in 1935) he would be awarded the Frederic Ives Medal. Obsessed with optical astronomy since childhood, Hale graduated from MIT in physics and studied solar physics at Harvard. Hale recognized the advantages of reflectors and in 1908 used a 60-inch-diameter glass disk given to him by his father to build the world's largest telescope on Mt. Wilson in southern California. By 1916 Hale had obtained funds from John D. Hooker, a Chicago philanthropist, and he was building, once again, the world's largest telescope: the 100-inch, dedicated in 1917. The 100-inch Hooker ground-based telescope is the same size as the Hubble Space Telescope of today. By 1935, Hale had sold the Rockefeller Foundation on supporting the design and construction of a 200-inch telescope and set off for a third time to build the world's largest telescope. George Ellery Hale engaged private financial support for optical telescopes from wealthy barons of the industrial revolution: Yerkes, Carnegie, Hooker, and Rockefeller. Figure 2 shows Hale with Andrew Carnegie in 1910. Hale established the tradition of private support that continues today with the Keck telescopes, Sloan Digital Sky Survey, and others.

Using new sensitive photographic emulsions developed by C. E. K. Mees (for whom the OSA Mees Medal is named), Edwin Hubble (shown in Fig. 3) imaged several Cepheid variables in the Andromeda Galaxy (M-31). The average luminosity of these variables is constant. Therefore, a measurement of the brightness of these very faint objects in M31 gives a direct measure of the distance. The measured distance was well outside our galaxy, demonstrating that spiral nebulae were outside our galaxy and thus proving that the universe was very large indeed! Hubble went on to show that the universe was expanding, thus providing fundamental evidence for today's "big bang" cosmology.

In 1930 an Estonian optician, Bernard Schmidt, developed his Schmidt camera for the imaging of large areas of the sky. For the first time, astronomers could make wide-FOV surveys needed to study the large-scale structure of our galaxy and to create catalogs of spectral types and variable stars in an efficient manner. The first large-aperture Schmidt cameras were the 40-cm-aperture at Mt. Palomar (1936) and the 60-cm at Case Western Reserve University (1939).

In 1946 Aden Meinel (1982 Ives Medalist, 1952 Lomb Medalist, and OSA President) built the first high-speed Schmidt camera and discovered the OH bands in the IR spectrum of the atmosphere using recently declassified infrared-sensitive photographic emulsions. James Baker (OSA Ives Medalist) improved on Schmidt's design to create the Baker–Nunn camera for wide-angle observations of artificial satellites passing rapidly overhead.

▲ **Fig. 1.** George Ellery Hale, astronomer and founding member of the OSA. Credit Huntington Library, San Marino, California. (The University of Chicago Yerkes Observatory, courtesy AIP Emilio Segre Visual Archives.)

Hale conceived the 200-inch telescope shortly after the dedication of the 100-inch telescope in 1917. The task of raising funds, keeping the vision alive, and preparing conceptual designs occupied most the 1920s. By 1928 Hale secured a grant of $6 million from the Rockefeller Foundation to complete the design and begin construction of the 200-inch telescope on Mt. Palomar. The Corning Glass Works, an OSA Corporate Member, working over a ten-year period, developed the technology and cast the Pyrex primary mirror. Construction of the observatory facilities began in 1936 but was interrupted by the onset of World War II. The telescope was completed and dedicated in 1948. Ira Bowen (1952 Ives Medalist) refined the optical system and the grating spectrographs and rebuilt the mirror support system. The telescope was not open for scientific use until 1949, and the first astronomer to use it was Edwin Hubble.

John Strong (1956 Ives Medalist and OSA President), demonstrated the advantages of using an evaporative aluminum coating on the 100-inch telescope in 1936. Before this, chemically deposited silver was used, which degraded rapidly to limit the faintest magnitude that could be recorded. The reflectivity of silver degrades significantly within a few days. Al coatings on mirrors are robust and with proper care retain high reflectivity for years. This increase in telescope transmittance enabled astronomers to record stars several magnitudes fainter than before.

During World War II, most optical astronomers were involved in the war effort. Scanners, detectors, photomultipliers, mirror coatings, manufacturing methods for large glass mirrors, and high-speed cameras were just a few of the technologies developed by optical astronomers during this period.

At the end of the war optical astronomers returned to civilian jobs. The new infrared-sensitive photographic films developed during the conflict were now used to extend astronomical discoveries into the infrared. Photomultipliers were used to make precision measurements of stellar brightness and color. These data improved our understanding of stellar evolution and reddening (absorption) due to interstellar matter.

The National Science Foundation was founded in 1950. Its earliest research center was the Kitt Peak National Observatory founded in 1955 operated under a board of directors from several university astronomy departments. Aden Meinel, an astronomy professor and optical scientist from the University of Chicago, was selected to be the founding director. The purpose of the observatory was to provide astronomical telescope time on a peer-review selection basis to all astronomers in the U.S. Under Meinel's direction the observatory developed the process for the thermal slump of a Pyrex mirror around a conformal mold (used in the 82-inch telescope), created a rocket program for UV spectroscopy of stellar objects, developed the world's largest solar observatory (the 60-inch McMath-Pierce), developed a 50-inch robot telescope for photoelectric photometry, and laid the groundwork for the first program in observational infrared astrophysics.

▲ **Fig. 2.** George Ellery Hale (right) possibly discussing future telescopes with Andrew Carnegie (center) in 1910. (Image courtesy The Observatories of the Carnegie Institution for Science Collection at the Huntington Library, San Marino, California.)

In 1960 Meinel left the Kitt Peak National Observatory to become the director of Steward Observatory. There he led the academic program, developed a 92-inch telescope for the University of Arizona on Kitt Peak Mountain and led an initiative to establish a national center of excellence in optical sciences and engineering, focused on many issues related to technology for astronomical telescopes and instruments. In 1964 funding became available, and the University of Arizona established the Optical Sciences Center under Aden's leadership. Aden established a distinguished faculty composed of A. F. Turner (Ives Medalist), R. R. Shannon (1985 OSA President), R. V. Shack (David Richardson Medalist), J. C. Wyant (2010 OSA President), and Roger Angel (OSA Fellow). Figure 4 shows Aden Meinel in 1985 while at NASA/JPL. In 1973 Aden resigned from the directorship to continue research in solar thermal energy, and Peter Franken (OSA Wood Prize and OSA President) became director.

In the late 1970s Roger Angel (OSA member) experimented with spin casting Pyrex mirrors for astronomical telescopes. This development has led to a family of 8-meter ground-based telescopes, which are revolutionizing our astrophysical understanding of the universe around the world.

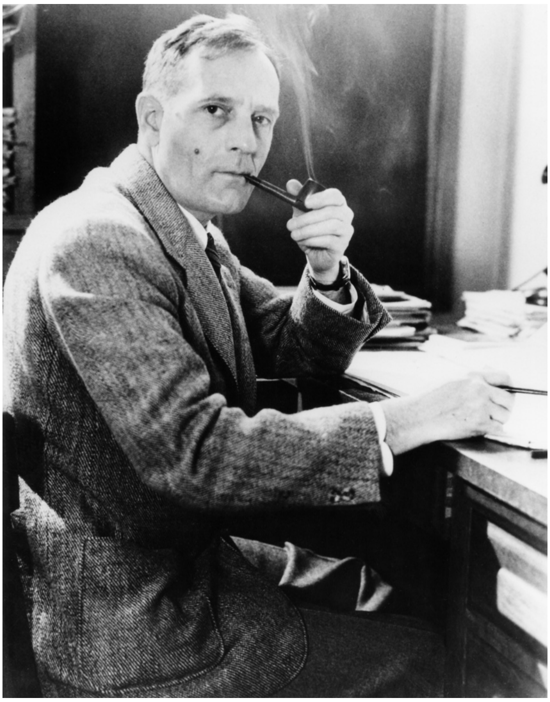
In 1920 optical physicist A. A. Michelson (OSA Honorary Member) made the first measurements of the diameter of a star using a white-light spatial interferometer mounted to the top of the 100-inch telescope. Atmospheric seeing and telescope stability prohibited useful data using the photographic plates of the time, and both he and his colleague F. G. Pease resorted to visual observations of flickering fringes to measure the diameter of stars. Breckinridge (OSA Fellow) recorded the first direct images of the fringes more than 50 years later. C. H. Townes (1996 Ives Medalist and Nobel

Laureate) developed the heterodyne-interferometer method and made early measurements of details of stellar atmospheres. Townes also invented the laser, which astronomers use in conjunction with adaptive optics to provide reference laser guide stars to remove atmospheric turbulence and enable diffraction-limited imaging from large-aperture ground-based astronomical telescopes. Over the past 30 years stellar optical interferometry has advanced to become a highly useful tool for the astronomy community. Today, several ground-based observatories use optical interferometry to measure high-angular-resolution (<0.001 arc sec) details across the surfaces of stars in the presence of Earth's atmospheric turbulence.

This 25-year period from 1975 to 2000 in the history of the OSA saw an explosive growth in technologies to make very large mirrors, long-baseline interferometers, large-area detectors, and space telescope systems. Angular resolution on the sky went from 0.5 arc sec to 0.001 arc sec and the surfaces of hundreds of stars were resolved. The high-speed electronics developed for military and commercial applications and innovative optical systems enabled long-baseline Michelson stellar interferometers for high-angular-resolution astronomy. Astronomers used atmospheric-turbulence-induced speckle patterns to create diffraction-limited images at large optical telescopes and thus make the first direct images across the surfaces of stars. The Orbiting Astronomical Observatory (OAO) was built and launched, and the Hubble Space Telescope (HST) was built and corrected.

Mt. Wilson astronomers discovered that larger telescopes, while collecting more photons than smaller telescopes, did not necessarily mean observing fainter objects. Atmospheric turbulence introduces wavefront errors as a function of time. Three major problems confronted the implementation of a system to correct atmospherically induced time-dependent phase perturbations. These were the need for (1) wavefront sensing, (2) a deformable mirror, and (3) signal and control processing.

Several OSA members pioneered practical solutions to these problems to increase the angular resolution on the sky from the seeing-limited 0.5 arc sec to 0.005 arc sec for a gain of 10,000 in area resolution. Although no one person was responsible for the invention of adaptive optics, OSA Fellows John Hardy and Mark Ealey and others from ITEK Optical Systems (OSA Corporate Member at the time) led the technology development of ground-based telescope systems to image distant objects



▲ **Fig. 3.** Edwin Hubble, who proved that the universe is much larger than we thought and is expanding. (Hale Observatories, courtesy AIP Emilio Segre Visual Archives.)

▲ **Fig. 4.** Aden Meinel in 1985 while at NASA/JPL. (Courtesy NASA/JPL-Caltech P-31041A.)

through atmospheric turbulence for the Air Force. At Kirtland Air Force Base Bob Fugate (OSA Fellow) demonstrated laser guide star adaptive optics, a technology in common use today at the Keck Telescope and a critical part of the new very large 30-meter-class telescopes. Figure 5 shows a laser guide star being used to compensate for atmospheric distortion.

▲ **Fig. 5.** A laser guide star tuned to the wavelength of sodium atoms in the atmosphere, providing information on atmospheric turbulence to allow for adaptive optics to compensate and enable improved telescope resolution. (©Laurie Hatch.)

Today there are four optical telescopes with apertures over 10 meters and nine 8-meter-class optical telescopes in operation nightly recording faint radiation from the cosmos. The Keck Ten-Meter-Diameter Telescope Project, under the technical leadership of Jerry Nelson (OSA Senior Member) pioneered the large aperture segmented phased telescope in common use today. OSA Corporate Members Corning Glass and Schott Glass and the University of Arizona under the leadership of OSA Fellow Roger Angel pioneered the design and cost-effective manufacture of monolithic mirror blanks 8 meters in diameter.

In 2016, on the occasion of the 100th anniversary of the OSA, there are three very ambitious projects underway to build astronomical optical telescopes with 30-meter-aperture-class phased primary mirrors. Each of these will be equipped with laser guide star adaptive optics to remove the effects of atmospheric turbulence and thus enable diffraction-limited imaging at resolutions approaching 3 milliarcsec steerable over a FOV on the order of 20 arc min. The Thirty Meter Telescope (TMT) will have over 500 phased mirror segments. The Giant Magellan Telescope (GMT) will have seven 8-meter mirrors in a hexagonal pattern with one of the mirrors at the center. The Extremely Large Telescope (ELT) will have 798 hexagonal segments each 1.45 meters across to create a 40-meter-diameter primary mirror.

The past 100 years of optical telescope development has led to profound changes in our understanding of the universe. The next 100 years of optical astronomy may reveal that mankind is not alone in the universe and that life exists and flourishes on planets around distant stars—stars so far away that our only contact will be with the optical photons reflected from the surface of exoplanets. Innovative spectrometers and polarimeters will be used to estimate the presence of life. Only if humans invent a way around the limits of speed-of-light travel will two-way communication with exoplanet life be possible.

1975–1990

Space Telescopes for Astronomy

James Breckinridge

In 1946, Lyman Spitzer of Princeton University proposed the construction of a space telescope for astrophysics, and Princeton astronomers launched several balloon-borne telescopes (Stratoscope project) to operate in the dry excellent seeing provided by the upper stratosphere to demonstrate the value of space science.

At the very beginning of NASA, Nancy Roman, Lyman Spitzer, and Art Code laid out a space satellite program that envisioned a series of modest-aperture telescopes for UV and optical astronomy [the Orbiting Astronomical Observatory (OAO)] and an R&D program leading to a "large space telescope." The seeds of the Hubble Space Telescope were sown 35 years before its launch.

In 1962 the world's first space telescope was launched, and it recorded the UV spectrum of the Sun. The OAO program became a series of three space telescopes. The first OAO was to carry experiments, and observing time was to be shared between the two university groups that produced the instruments. However, when that satellite was launched, it almost immediately self-destructed before the scientific instruments could be turned on.

NASA quickly organized an additional launch using flight spares of the satellite and the scientific instruments. That satellite was successful and is referred to now as OAO-2. It was launched 7 December 1968, carried 11 UV telescopes, and operated until 1973. OAO-2 discovered that comets are surrounded by enormous halos of hydrogen several hundred thousand kilometers across and made observations of novae to find that their UV brightness often increased during the decline in their optical brightness.

OAO-3 (Copernicus) was orbited in August of 1972 and carried an 80-cm-diameter telescope for UV astronomy. OAO-3 successfully operated for 14 years and established an excellent reputation for the highest-quality astronomical data at the time. The Copernicus mission played a large role in winning the support of the wider astronomical community for space astronomy, not only because of the very high-quality data it produced, covering the UV to below the Lyman limit, but also because of the serious commitment Spitzer and his Princeton colleagues showed to making the data available and easily interpretable. Complete spectra were obtained for only about 500 stars, very modest by today's standards. But the scientific impact of those spectra was huge!

The concept for a series of four large telescopes, called the "Great Observatories," evolved at NASA starting in the 1980s. In order of increasing wavelength they were Compton Gamma Ray Observatory (CGRO), Advanced X-Ray Astrophysics Facility (AXAF), now called Chandra, the Hubble Space Telescope (HST), and the Space Infrared Telescope Facility (SIRTF), now called Spitzer. Optical Society (OSA) members had a major role in the development of AXAF, Chandra, and Spitzer.

HST started out as the Large Space Telescope (LST) with a 3-meter aperture. Soon the reality of the launch vehicle capacity set in and NASA issued a request for information to the industry for a 2.4-meter-diameter telescope. Three optics companies, all corporate members of OSA, responded with feasibility studies: Eastman Kodak, Itek, and Perkin-Elmer. Perkin-Elmer was selected as the primary telescope provider. NASA recognized that the longest lead item in the procurement would be the primary mirror and directed Perkin-Elmer to fund Eastman Kodak to provide a back-up mirror. This mirror is now at the Smithsonian Air and Space Museum. Corning manufactured both of the ultra-low expansion (ULE) honeycomb 2.4-meter mirror blanks. PE was responsible for the telescope, and Lockheed Sunnyvale was the spacecraft system integrator.

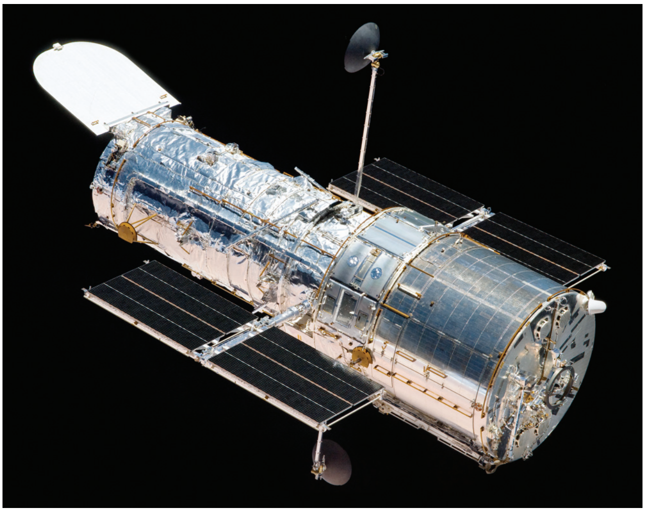
"Large" was dropped from the LST name during its development, and later it was renamed after Edwin Hubble to become the HST. NASA Headquarters issued a competitive-science solicitation for instruments. These UV/optical/IR science instruments were designed to be replaced on-orbit.

The HST became the world's first scientific instrument with the capability to be serviced multiple times on-orbit. The instruments selected were the Wide-Field Planetary Camera (WF/PC), the Faint Object Camera (FOC), the Faint Object Spectrograph (FOS), the Goddard High Resolution Spectrograph (GHRS), and the High Speed Photometer (HSP). The HST primary mirror was maintained near room temperature. That combined with the poor IR detectors at the time prohibited an infrared astronomy instrument.

HST was scheduled for launch in 1986 soon after the Challenger mission that ended in disaster. The shuttle fleet was grounded for 32 months, delaying the HST launch to late April 1990. By the end of May 1990 it was discovered that the telescope could not be focused, and in June the error was suggested to be spherical aberration. NASA headquarters formed two teams. One, the official NASA optical failure review board led by Dr. Lew Allen (a retired four-star general and JPL director) had membership and support from Optical Society Fellows Roger Angel, Bob Shannon, John Mangus, Jim Breckinridge, and Bob Parks. This team investigated the root cause of the error. The other board, the Hubble Independent Optical Review Panel (HIORP) was led by Optical Society Fellow Duncan Moore. Optical Society Fellows Aden and Marjorie Meinel, Dietrich Korsch, Dan Schulte, Art Vaughan, and George Lawrence, among others, were members. The HIORP had broad membership from the optics and astronomy communities and was charged with making recommendations on how to fix the error. The nation's optics community came together to establish that the error was on the primary. Nine optics groups composed of many Optical Society Members and Fellows across the country made independent measurements on the PE test apparatus hardware and on digital images recorded by the hardware on-orbit. The recording of star images across the field of view and at different telescope focus settings provided a diverse set of image data for the new prescription retrieval algorithms. For the first time, the on-orbit optical prescription was determined precisely. The intensity of this work is evidenced by the fact that it was completed over a ten-week period to meet the instrument rebuild schedule for a repair mission launch.

An accurate value for the telescope primary-mirror conic constant and the fact that the error was isolated to the primary enabled corrective optics to be integrated into a newly built WF/PC2 (designed and built by NASA/JPL), and a new optical system called COSTAR. COSTAR was designed and built by Ball, an Optical Society Corporate Member. Both instruments were inserted into HST on the first repair mission. The COSTAR optical system replaced the HSP instrument. This new optical system corrected the wavefront for the Faint Object Spectrograph (FOS), the Faint Object Camera (FOC), and the GHRS. In 1997 the IR system NICMOS was launched replacing COSTAR to give the telescope its first IR capability to 2-μm wavelength.

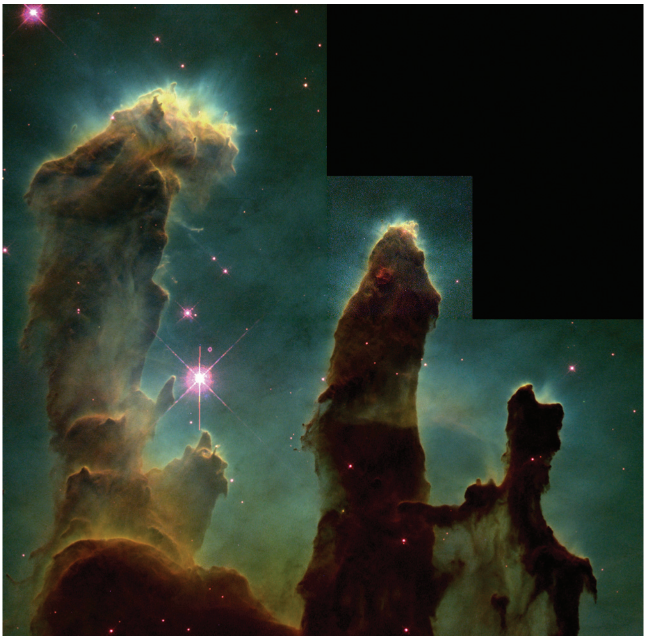
Today, at the one hundredth anniversary of The Optical Society, the HST has been successfully operating for 26 years. By far it is the most productive scientific UV/optical instrument ever known, a spectacular monument to the space optics community and the many dedicated Optical Society members who saved the mission from disaster. Figure 1 is a photo of the HST in orbit taken from the space shuttle after a service mission. One of the most famous and spectacular photos taken by HST is shown in Fig. 2. It is the so-called "pillars of creation" in the Eagle Nebula where stars, and by implication their



▲ **Fig. 1.** The HST in orbit. (Image courtesy of NASA.)

exoplanet systems, are seen forming in the dust clouds.

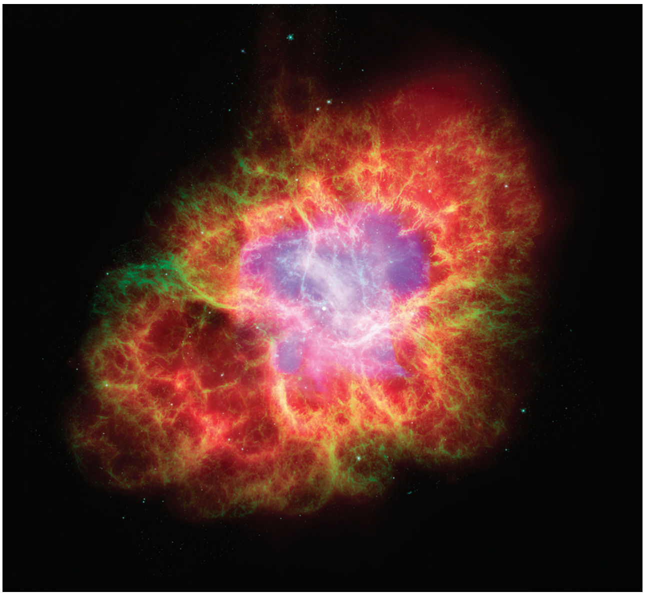
The x-ray telescope mission, Chandra, was launched in 1999, 33 years after the proposal by Riccardo Giacconi and Harvey Tananbaum. Chandra uses two sets of nested-cylinder mirrors in the hyperbola–parabola configuration of the Woljter type-2-configuration grazing-incidence x-ray telescope built by Eastman Kodak. Chandra's angular resolution is unmatched: between 80% and 95% of the incoming x-ray energy is focused into a 1-arcsec circle. Leon van Speybroek led the details of the optical design and the fabrication of the mirrors. Furthermore, x-rays reflect only at glancing angles, like skipping pebbles across a pond, so the mirrors must be shaped like cylinders rather than the familiar dish shape of mirrors on optical telescopes. The Chandra X-ray Observatory contains four co-aligned pairs of mirrors. Figure 3 shows an image of the Crab Nebula recorded with the ACIS instrument superposed upon an image recorded with HST to show the value of multispectral (visible and x-ray) imaging science.



▲ **Fig. 2.** The "pillars of creation." Star formation in the Eagle nebula photographed by the HST. (Image courtesy of NASA, ESA, STScI, J. Hester and P. Scowen [Arizona State University].)

Today, at the one hundredth anniversary of Optical Society, the Chandra has been successfully operating for 15 years, three times its design lifetime, and it remains in highly productive operation.

Much excellent IR astronomy from telescopes on the ground has been done through those spectral windows in the IR not absorbed by the Earth's atmosphere. However, many exciting astrophysics problems require the measurements of cold gas and dust available only using IR space telescopes, which measure the temperature of the universe and need to be colder than the sky they measure. Two major space cryogenic IR telescopes were designed, built, and launched to map the IR sky: the Infrared



▶ **Fig. 3.** This composite image uses data from three of NASA's Great Observatories. The Chandra x-ray image is shown in light blue, the HST optical images are in green and dark blue, and the Spitzer Space Telescope's infrared image is in red. The size of the x-ray image is smaller than the others because ultra-high-energy x-ray-emitting electrons radiate away their energy more quickly than the lower-energy electrons emitting optical and infrared light. The neutron star, which has mass equivalent to the Sun crammed into a rapidly spinning ball of neutrons 12 miles across, is the bright white dot in the center of the image. (X-Ray: NASA/CXC/J. Hester [ASU]; Optical: NASA/ESA/J. Hester & A. Loll [ASU]; Infrared: NASA/JPL-Caltech/R. Gehrz [Univ. Minn.])

Astronomical Satellite (IRAS) and Spitzer. Launched in 1983, the IRAS telescope system, whose scientific development was led by Gerry Neugebauer, was the first space observatory to perform an all-sky survey at IR wavelengths. Engineering and development of the optical system was completed at Ball Aerospace, an Optical Society Corporate Member, teamed with Steve Macenka, an Optical Society Fellow, at JPL. IRAS discovered over 350,000 new sources, including stellar gas and dust envelopes now known to be the birthplaces of exoplanet systems, some possibly similar to our own solar system. The Spitzer telescope system, the fourth and final telescope in the Great Observatory series, was launched in 2003 into an Earth-trailing orbit. The primary, secondary, and metering structure are all fabricated from berylium. The optics were configured at Tinsley, and cryo testing was carried out at JPL. Diffraction-limited imaging at 6.5 μm over a 30-arc-min field of view was achieved.

By the year 2000, plans were underway to build an even larger space telescope, and NASA funded the Next Generation Space Telescope (NGST) study, which led to the James Webb Space Telescope (JWST), now scheduled for launch in 2019, in time to start the second hundred years of The Optical Society. John Mather, Nobel Prize Laureate in Physics 2006 and Optical Society Fellow, was the chief scientist for the project during its formative years. This telescope builds on the success of the large ground-based segmented telescopes, e.g., Keck. Telescopes with segmented primary mirrors that are mechanically deployed once the spacecraft is in orbit make possible very large space telescopes.

Recently several smaller space optics systems have revolutionized our understanding of the universe. These are: WISE, COBE, GALEX, Herschel, Planck, WIRE, WISE, and WMAP. The SOFIA is a 3-meter telescope mounted in a B747 for IR observations above the atmosphere. The Kepler space telescope launched March 2009 is a 0.95-meter clear-aperture Schmidt camera precision radiometer that contains arrays of CCDs totaling 95 megapixels staring at 140,000 stars across a FOV of 105 deg^2 in the constellation of Cygnus. The Kepler mission has discovered several thousand exoplanets and will continue to revolutionize our understanding of the evolution of planetary systems, stellar atmospheres, and stellar interiors as the enormous database is analyzed in detail over the coming decade.

Today one of the most exciting space optics programs is the design and construction of hyper-contrast optical systems to characterize exoplanets in the presence of the intense radiation from the central star of the exoplanet system. Terrestrial planets are 1 part per trillion as bright as the central star. Spectrometric measurements are required of the radiation reflecting and emitting from the exoplanet. These measurements provide data to estimate planetary surface and atmospheric composition! Direct observation of rocky terrestrial planets, which might harbor life as we know it, requires large-aperture telescopes. This is an opportunity to answer one of humanity's most compelling questions: Are we alone in the universe?

In addition to using spectrometric measurements to resolve the question of composition, optical spectrometers are also used to determine the radial (along the line of sight) velocity as a function of time to an accuracy of centimeters per second. Precision optical astrometry is used to determine the motion of stars across the sky to precisions approaching microarcsec. These two measurements provide the data we need to calculate the orbit of the planet about its parent star.

Direct images and spectra of exoplanets at contrast levels of 10^{-10} are needed so astronomers can record the light reflected from the exoplanet and search for life signatures in the atmosphere and on the surface. All of these require new-technology optical systems operating in the harsh space environment out from under the turbulence of the Earth's atmosphere. Today, astronomical science, enabled by innovative optical telescope and instrument design, is on the threshold of revealing details on the evolution of the universe and the presence of life beyond Earth.

The JWST is the largest space optical system under construction now. It represents the state of the art in optical design, engineering, fabrication, and testing. The JWST will replace the spectacularly successful Hubble Space Telescope with a much more capable system promising further astounding discoveries.

1975–1990

Contact Lenses for Vision Correction: A Journey from Rare to Commonplace

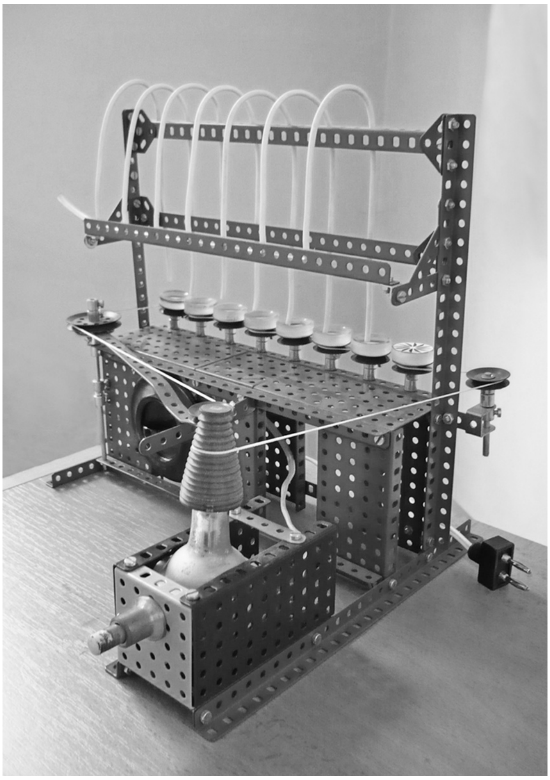
Ian Cox

Although the first practical contact lens was described in 1888 [1], glass-blown shells formed individually to rest on the sclera and vault across the cornea were the norm until the 1930s. The advent of polymethyl methacrylate (PMMA) made it possible, in a method pioneered by William Feinbloom [2], to process an all-plastic lens that could be fitted by custom molding or trial fitting from a range of premade lenses. This reduced the weight and cost of lenses while improving comfort and wearing times. It was not until 1948 that Kevin Tuohy, an optician, made the first corneal contact lens [3]. Accidentally cutting through a scleral shell at the edge of the optic zone, Tuohy tried the small-diameter lens that was left on his own eye and quickly realized that a lens fitted within the cornea could be more comfortable and provide longer wearing times than a scleral shell. The realization by Smelser and Ozanics that oxygen for corneal metabolism came directly from the atmosphere led to a major shift to corneal contact lenses because the fit could be adjusted to replenish the oxygenated tear film with every blink, thus extending comfortable wearing times from just a few hours. The contact lens market expanded with commercially available corneal contact lens designs enabling the correction of myopia, hyperopia, astigmatism, and even novel bifocal designs for presbyopia correction.

Otto Wichterle (Fig. 1) was a brilliant Czech polymer chemist who made the world's first "soft" contact lenses from his newly invented HEMA hydrogel material [4]. This 38% water content material was highly flexible, oxygen permeable, and significantly more comfortable than the rigid PMMA corneal contact lenses that were available. Although working behind the "Iron Curtain," an American patent company acquired the intellectual property rights from Wichterle and licensed them to Bausch & Lomb (B&L). The company licensed both the material and the novel "spincasting" manufacturing technique that Wichterle had developed in his own kitchen. The prototype for this production method was built from an erector set, powered by the electric motor from his phonograph (Fig. 2). Henry Knoll, a physicist working at B&L and one of a team assigned to developing the Wichterle prototypes, pointed out the difficulty in working with this hydrogel material. "The first lens we released commercially was called the C series lens, we built the A series and the B series but neither would stay on the eye after a few blinks. Management said if the third design didn't work we would give up on the project." The C-series contact lens design (Fig. 3) fitted the eye, and although the optics were compromised by the wildy aspheric posterior lens surface produced by the "spincasting" manufacturing process, the lens was a commercial success when launched in 1971 following FDA approval. The dramatically improved comfort changed the contact lens industry in the U.S., and ultimately the world, with rigid corneal contact lenses today accounting for less than 10% of the lenses fitted worldwide. Otto Wichterle was recognized for his great contributions to the world of optics when he was awarded the R. W. Wood Medal by The Optical Society (OSA) in 1984.

Initially available only in spherical powers to correct myopia and later hyperopia, soft lenses to correct astigmatism were first introduced in the U.S. in the early 1980s. Unlike rigid lenses which "mask" the astigmatic component of the cornea, soft lenses conform to the

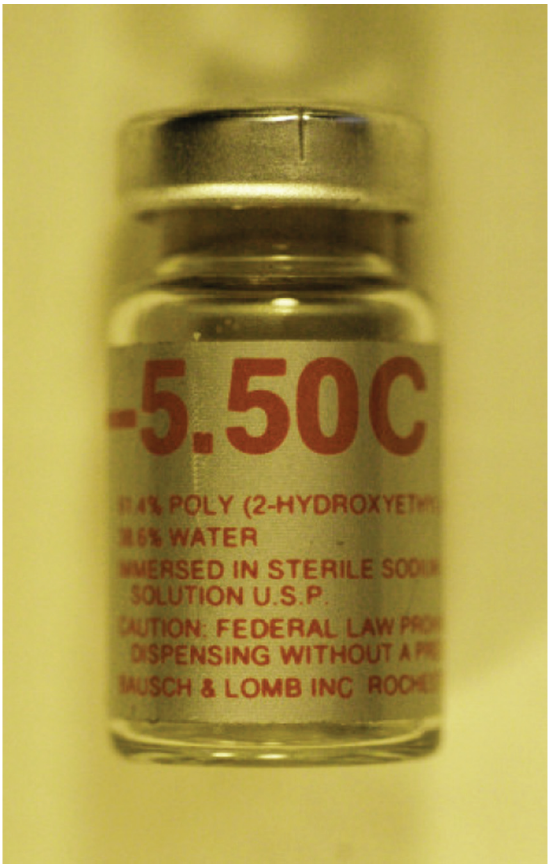
▲ **Fig. 1.** Otto Wichterle, Czech polymer chemist, inventor of the first hydrogel material to be used in making soft contact lenses. Wichterle was responsible for making the first usable soft contact lenses in his lab behind the "Iron Curtain." (AIP Emilio Serge Visual Archives, Physics Today Collection.)



▲ **Fig. 2.** A model of the first spin-casting machine that Otto Wichterle used to make the first soft contact lenses in his kitchen.

underlying corneal shape, requiring a method of stabilization and orientation to be built into the physical shape of the lens. The most successful designs used an increasing thickness profile in the vertical meridian of the lens, allowing the squeeze force of the upper eyelid to stabilize the lens on the eye between blinks. Multifocal soft lenses designed to correct presbyopia were introduced by B&L and CIBA VISION in 1982. B&L used its early experience with significant spherical aberration in its first lenses for myopia to help manufacture a lens with sufficient spherical aberration to expand the depth of field of the wearer. Ironically, after spending years trying to eliminate spherical aberration inherent in the "spincast" lens product, B&L was purposely designing it in the lens with the PA1 bifocal.

A major issue with soft contact lenses over the 1970s and 1980s was combating adverse ocular responses related to deposition of protein and lipid on lens surfaces from the tear film. This required daily cleaning and disinfection routines and impacted the longevity of the lenses, prescribed as a single pair to be worn daily for as long as they lasted, typically a year or more. A second issue was transmitting sufficient oxygen from the atmosphere through lenses to ensure an adequate physiological environment for the cornea. Many patients had their lens wear curtailed from insufficient oxygen being available to the eyes during wearing. This was also the time of "continuous wear," a modality where patients wore their contact lenses constantly, with removal as needed for cleaning (typically every 30 days in the early 1980s) [5]. Although convenient, continuous wear only exacerbated the issues of deposition, reduced lens life, and caused a significant increase in ocular adverse responses due to reduced oxygen availability to the cornea. In 1982, a small company in Denmark started cast molding hydrogel contact lenses and packaging them in small plastic blisters with foil covers. All other companies delivered their lenses individually, stored in a small glass serum vial, packaging that dated back to the original B&L lens. Danalens was the first "disposable" contact lens and lit the fuse on a major upheaval in the contact lens industry (Fig. 4). Johnson and Johnson, sensing an opportunity to enter the lucrative contact lens market in the U.S., acquired the Danalens production process and a small contact lens company called Vistakon whose hydrogel lens material was already approved by the FDA. Within five years Vistakon launched the first disposable lens in the United States (1987). Launched as a continuous-wear lens to be replaced weekly, the marketplace eventually dictated its use as a daily wear only (no overnight wear) lens with a biweekly replacement schedule. Although the oxygen permeability of these new lenses was no better, the fact that patients could buy them for only a few dollars each (previously patients would typically pay hundreds of dollars for a pair of lenses) and replace them frequently made them a rapid success. Toric and multifocal options soon followed as companies invested in the manufacturing capacity necessary to process these complex designs for a low cost. As manufacturing technology improved and cost of goods decreased, the option of a truly disposable lens, one that was worn once and then discarded, became a reality. Vistakon again led the industry by launching the first daily disposable contact lens in 1994. Although the cost of each lens was less than one dollar to the patient, the high annual cost prohibited rapid adoption of daily disposables, and it was another decade before this modality made any significant inroads into the marketplace.



▲ Fig. 3. The first commercially available soft lens in the U.S., the Bausch & Lomb “C” Series. (Courtesy of Andrew Gasson.)

In the intervening years, others were still chasing the ultimate in convenience, a lens that was so physiologically compatible with the eye that it could be worn continuously for 30 days without the risk of adverse ocular responses. The massive oxygen permeability of silicone elastomer led researchers to develop lenses made from this material in the late 1970s, with Dow Corning being the most well-known manufacturer to try this alternative material. Although physiologically successful, silicone elastomer lenses had one undesirable and potentially dangerous flaw: their rubber-like nature generated negative pressure under the lens during wear and resulted in the lens sticking to the eye. The only path forward was a hybrid material, a silicone hydrogel. Although seemingly simple, material scientists were essentially trying to mix “oil and water” and maintain a transparent material. B&L, the first company to bring soft hydrogel contact lenses to the market in 1972, were also the first to develop a commercially viable silicone hydrogel lens. This lens provided four times the oxygen transmission of hydrogel lenses, and it was approved for up to 30 days of continuous wear in 1999. Clinicians immediately noted that highly oxygen transmissive lenses eradicated significant adverse responses related to oxygen deprivation at the cornea, but they were slow to adopt silicone hydrogel lenses due to the up to 30 days continuous-wear indication awarded by the FDA. Experience over the years had shown that corneal ulcers, or microbial keratitis, was the single most significant adverse response associated with continuous wear, with the FDA limiting approval of all hydrogel lenses to six nights maximum in 1989 over their concern with incidence levels. Clinicians and companies now recommend silicone hydrogel lenses for daily wear or extended wear with monthly or more-frequent replacement, but the largest area of growth within the contact lens industry is the daily wear modality.

Currently available soft lens materials provide excellent physiological compatibility with the eye, and the cornea specifically, when worn in a daily wear modality, and so the focus of the industry has

▲ **Fig. 4.** Examples of the first disposable soft contact lens. Although the Danalens lens design and material were not unique, the packaging and delivery concept were innovative and ultimately changed the way contact lenses were sold the world over. (Courtesy of Andrew Gasson.)

moved to improving end-of-day comfort through design and material formulation, as well as improved optical performance. This last development has been driven by the development of clinically applicable Hartmann–Schack wavefront sensors. Porter *et al.* [5] measured the wavefront error of the eye of a large contact-lens-wearing population, identifying that the Strehl ratio of the eye can be significantly improved by correcting at least the major higher-order wavefront aberrations. This technique proved to be an ideal method to evaluate the optical performance of contact lenses on and off the eye, and OSA members led the development of standards for reporting the optical aberrations of eyes. Ideally, individual prescription contact lenses should be made for each eye based on wavefront measurements performed in a clinical setting, enabling correction of all higher-order aberrations for improved low-light vision. Although the feasibility of this concept has been demonstrated by Marsack, the challenge for industry is to deliver these custom-optics contact lenses in the same low-cost, disposable paradigm that patients and clinicians are currently using. In the meantime, at least one manufacturer (B&L) is altering the inherent spherical aberration of their spherical and toric contact lens products using aspheric optical surfaces to minimize the spherical aberration magnitude of the eye with the lens in place and improve the quality of vision under low-illumination conditions.

References

1. R. M. Pearson and N. Efron, “Hundredth anniversary of August Müller’s inaugural dissertation on contact lenses,” Surv. Ophthalmol. **34**, 133–141 (1989).
2. R. B. Mandell, *Contact Lens Practice*, 4th ed. (Charles C. Thomas, 1988).
3. K. M. Tuohy, “Contact lens,” U.S. patent 2,510,438, filed 28 February 1948. Issued 6 June 1950.
4. O. Wichterle and D. Lim, “Hydrophilic gels for biological use,” Nature **185**, 117–118 (1960).
5. J. Porter, A. Guirao, I. G. Cox, and D. R. Williams, “Monochromatic aberrations of the human eye in a large population,” J. Opt. Soc. Am. A **18**, 1793–1803 (2001).

Excimer Laser Surgery: Laying the Foundation for Laser Refractive Surgery

James J. Wynne

Discovery of Excimer Laser Surgery

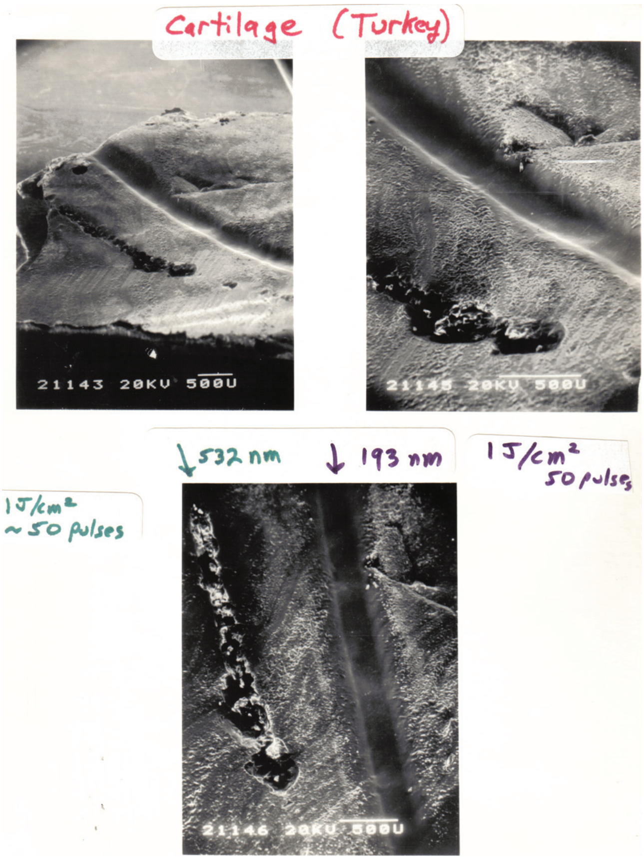
On 27 November 1981, the day after Thanksgiving, Rangaswamy Srinivasan brought Thanksgiving leftovers into the IBM Thomas J. Watson Research Center, where he irradiated turkey cartilage with ~10-ns pulses of light from an argon fluoride (ArF) excimer laser. This irradiation produced a clean-looking "incision," as observed through an optical microscope. Subsequently, Srinivasan and his IBM colleague, Samuel E. Blum, carried out further irradiation of cartilage samples. Srinivasan gave a sample to the author, and, for comparison, it was irradiated with ~10-ns pulses of 532-nm light from a Q-switched, frequency-doubled, Nd:YAG laser. This irradiation did not incise the sample; rather it created a burned, charred region of tissue. Figure 1 shows three different views and magnifications of scanning electron micrographs (SEMs) of the sample, revealing the stunningly different morphology of the two irradiated regions: the clean incision with no evidence of thermal damage, etched steadily deeper by a sequence of pulses of 193-nm light, and the damaged region produced by the pulses of 532-nm light.

Realizing that Srinivasan, Blum, and the author had discovered something novel and unexpected, they wrote an invention disclosure, describing multiple potential surgical applications. They anticipated that the absence of collateral damage to the tissue underlying and adjacent to the incision produced in vitro would result in minimal collateral damage when the technique was applied in vivo. The ensuing healing would not produce scar tissue. This insight, a radical departure from all other laser surgery, was unprecedented and underlies the subsequent application of their discovery to laser refractive surgery.

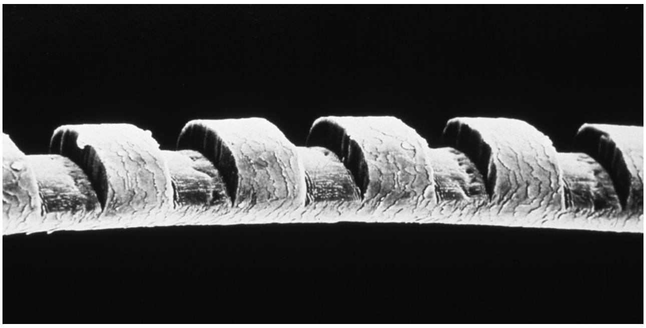
Background to This Discovery

As manager of the Laser Physics and Chemistry department at the Watson Research Center, one of the author's responsibilities was to ensure that there was access to the best and latest laser instrumentation. When the excimer laser became commercially available, the author purchased one for use by the scientists in his department. Since 1960, Srinivasan had been studying the action of ultraviolet radiation on organic materials, e.g., polymers. In 1980, he and his technical assistant, Veronica Mayne-Banton, discovered that the ~10-ns pulses of far ultraviolet radiation from the excimer laser could photo-etch solid organic polymers, if the fluence of the radiation exceeded an ablation threshold [1,2].

Srinivasan and the author then speculated about whether an animal's structural protein, such as collagen, which contains the peptide bond as the repeating unit along the chain, would also respond to the ultraviolet laser pulses. They knew that when skin was incised with a sharp blade, the wound would heal without fibrosis and, hence, no scar tissue. Conceivably, living skin



▲ Fig. 1. Three scanning electron micrographs of laser-irradiated turkey cartilage, recorded from different perspectives and with different magnification. In the bottom micrograph, arrows indicate the regions irradiated with 193-nm light and 532-nm light. For each wavelength, the fluence/pulse and number of pulses of irradiation are given.



▲ Fig. 2. Scanning electron micrograph of a human hair etched by irradiation with an ArF excimer laser; the notches are 50 μm wide.

or other tissue, when incised by irradiation from a pulsed ultraviolet light source, would also heal without fibrosis and scarring.

Physics of Ablation

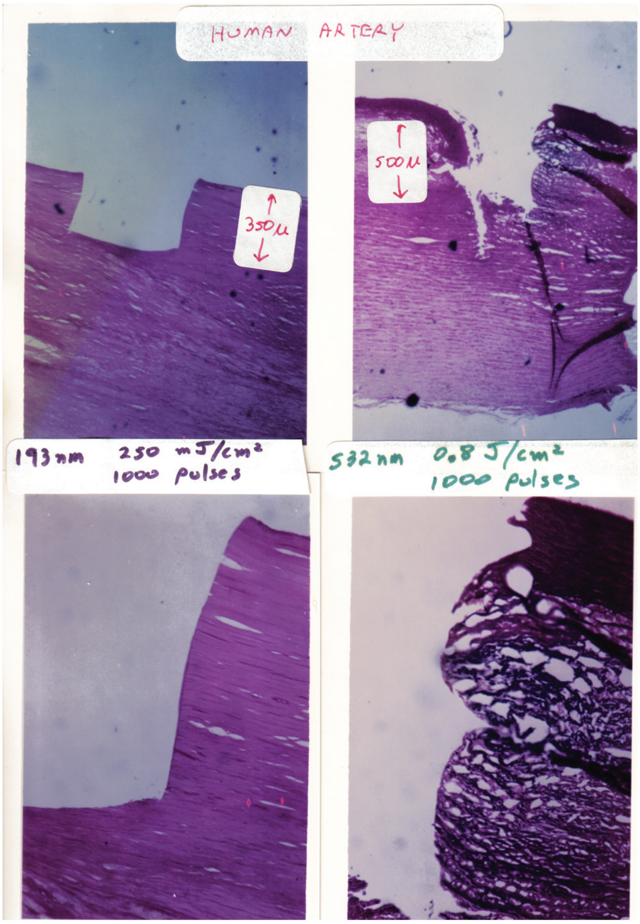
Ablation occurs when the laser fluence is such that the energy deposited in a volume of tissue is sufficient to break the chemical and physical bonds holding the tissue together producing a gas that is under high pressure. The gas then expands away from the irradiated surface, carrying with it most of the energy that was deposited into the volume that absorbed the energy. If the absorption depth is sufficiently shallow and the pulse duration is sufficiently short, the expanding gas can escape from the surface in a time that is short compared with thermal diffusion times, leaving a clean incision with minimal collateral damage. These conditions are readily satisfied by a short pulse of short-wavelength light having sufficient energy/unit area, given that protein and lipids are very strong absorbers of ultraviolet light.

Next Steps

To develop practical innovative applications, Srinivasan, Blum, and the author needed to collaborate with medical/surgical professionals. To interest these professionals, they etched a single human hair by a succession of 193-nm ArF excimer laser pulses, producing an SEM micrograph (Fig. 2), showing 50-μm-wide laser-etched notches.

While IBM was preparing a patent application, Srinivasan, Blum, and the author were constrained from discussing their discovery with people outside IBM. But a newly hired IBM colleague, Ralph Linkser, with an M.D. and a Ph.D. in physics, obtained fresh arterial tissue from a cadaver, and Linsker, Srinivasan, Blum, and the author irradiated a segment of aorta with both 193-nm light from the ArF excimer laser and 532-nm light from the Q-switched, frequency-doubled Nd:YAG laser. Once again the morphology of the tissue adjacent to the irradiated/incised regions, examined by standard tissue pathology techniques (Fig. 3), was stunningly different, with irradiation by the 193-nm light showing no evidence of thermal damage to the underlying and adjacent tissue [3].

This experimental study on freshly excised human tissue confirmed that excimer laser surgery removed tissue by a fundamentally new process. Srinivasan, Blum, and the author's vision—that excimer laser surgery would allow tissue to be incised so cleanly that subsequent healing would not produce scar tissue—was more than plausible; it was likely, subject to experimental verification on live animals.



▲ **Fig. 3.** Left side: Photo micrographs of human aorta irradiated by 1000 pulses of ArF excimer laser 193-nm light; lower image is a magnified view of the right-hand side of the laser-irradiated region. Right side: Photo micrographs of human aorta irradiated by 1000 pulses of Q-switched, frequency-doubled Nd:YAG laser 532-nm light; lower image is a magnified view of the right-hand side of the laser-irradiated region. (By permission of John Wiley & Sons, Inc.)

First Public Disclosure

After their patent application was filed, Srinivasan, Blum, and the author submitted a paper to *Science* magazine. Their paper was rejected because one of the referees argued that irradiation with far-ultraviolet radiation (far-UV) would be carcinogenic, making the technique more harmful than beneficial. Since Srinivasan had been invited to speak about his work on polymers at the upcoming CLEO 1983 conference co-sponsored by the OSA, Srinivasan, Blum, and the author wanted to get a publication into print as soon as possible. Therefore, they resubmitted their paper to *Laser Focus*, including some remarks about the new experiments on human aorta, and the *Laser Focus* issue containing their paper [4] was published simultaneously with CLEO 1983. Srinivasan's talk on 20 May, entitled "Ablative photodecomposition of organic polymer films by far-UV excimer laser radiation," included the first public disclosure that the excimer laser cleanly ablated biological specimens, as well as organic polymers.

From Excimer Laser Surgery to ArF Excimer Laser-based Refractive Surgery

At that very same CLEO 1983 meeting, Stephen Trokel and Francis L'Esperance, two renowned ophthalmologists, gave invited talks on applications of infrared lasers to ophthalmic surgery. The author attended both of their talks and was amazed at the results they obtained in successfully treating two very different ophthalmic conditions that were not candidates for excimer laser treatment. However, Trokel knew of ophthalmic conditions, such as myopia, that could be corrected by modifying the corneal curvature. A treatment known as radial keratotomy (RK) corrected myopia by using a cold steel scalpel to make radial incisions at the periphery of the cornea. Upon healing, the curvature of the front surface of the cornea was reduced, thereby reducing myopia. While this technique rarely yielded uncorrected visual acuity of 20/20, the patient's myopia was definitely reduced. One serious drawback of RK was that the depth of the radial incisions left the cornea mechanically less robust. The healed eye was more susceptible to "fracture" under impact, such as might occur during an automobile collision. Trokel speculated that the excimer laser might be a better scalpel for creating the RK incisions.

Upon learning of Srinivasan, Blum, and the author's discovery of excimer laser surgery, Trokel, who was affiliated with Columbia University's Harkness Eye Center in New York City, contacted Srinivasan and brought enucleated calf eyes (derived from slaughter) to the Watson Research Center on 20 July 1983. Srinivasan's technical assistant, Bodil Braren, participated in an experiment using the ArF excimer laser to precisely etch the corneal epithelial layer and stroma of these calf eyes. The published report of this study is routinely referred to by the ophthalmic community as the seminal paper in laser refractive surgery [5].

To conduct studies on live animals, the experiments were moved to Columbia's laboratories. Such experiments were necessary to convince the medical community that living cornea etched by the ArF excimer laser does not form scar tissue at the newly created surface and the etched volume is not filled in by new growth. The first experiment on a live rabbit in November 1983 showed excellent results in that, after a week of observation, the cornea was not only free from any scar tissue but the depression had not filled in. Further histological examination of the etched surface at high magnification showed an interface free from detectable damage.

L'Esperance, also affiliated with Columbia, thought beyond RK and filed a patent application describing the use of excimer laser ablation to modify the curvature of the cornea by selectively removing tissue from the front surface, not the periphery of the cornea. His U.S. patent 4,665,913 [6] specifically describes this process, which was later named photorefractive keratectomy (PRK).

Soon ophthalmologists around the world, who knew of the remarkable healing properties of the cornea, were at work exploring different ways to use to excimer lasers to reshape the cornea. From live animal experiments, they moved to enucleated human eyes, then to blind eyes of volunteers, where they could study the healing. Finally, in 1988, a sighted human was treated with PRK and, after the cornea had healed by epithelialization, this patient's myopia was corrected.

Development of an alternative technique, known as laser in situ keratomileusis (LASIK) commenced in 1987. In LASIK, a separate tool is used to create a hinged flap at the front of the cornea, preserving the epithelial layer and exposing underlying stroma, which is then irradiated and reshaped by the ArF excimer laser. After such irradiation, the flap is repositioned over the irradiated area, it adheres rather quickly, and the patient is soon permitted to blink, while the surgeon makes sure that the flap stays in place. No sutures are required. The flap acts like the cornea's own "bandaid," minimizing the discomfort of blinking. LASIK offers the patient much less discomfort than PRK and much more rapid attainment of ultimate visual acuity following surgery. For these reasons patients prefer LASIK to PRK, and far more LASIK procedures are performed than PRK procedures.

However, patients whose corneas are much thinner than average are not good candidates for LASIK, because a post-LASIK cornea is mechanically weaker than a post-PRK cornea, making the cornea more susceptible to impact or high-acceleration injury. In fact, the U.S. Navy accepts candidates into training programs for the Naval Air Force who had their visual acuity improved by PRK, but it does not accept candidates who had LASIK.

Pervasiveness of Laser Refractive Surgery

Since the U.S. Food and Drug Administration (FDA) granted approval to manufacturers of laser refractive surgery systems in 1995, more than 30 million patients have undergone the procedure to improve their eyesight. While patients choose to undergo this procedure for the obvious cosmetic reasons, many patients are unable to comfortably wear contact lenses. PRK and LASIK offer them a safe alternative that actually may cost less than the accumulated cost of wearing and maintaining contact lenses. Further, the U.S. military encourages its ground troops to have laser refractive surgery to eliminate the problems inherent in wearing glasses or contact lenses in combat situations (e.g., the desert sands of the Middle East). Laser refractive surgery can restore visual acuity to better than 20/20 as is required for certain aviators. With further refinements in so-called "custom wavefront-guided" laser refractive surgery, soon there may be a time when patients undergoing laser refractive surgery may expect to achieve visual acuity of 20/10.

Public awareness and interest in laser eye surgery was intense even before FDA approval. On 30 January 1987, *The Wall Street Journal* published an article entitled "Laser shaping of cornea shows promise at correcting eyesight," and on 29 September 1988, *The New York Times* published its first article on PRK, entitled "Laser may one day avert the need for eyeglasses." Subsequent articles in the press dealt with the progress in the research on PRK, the formation of three U.S. companies to market this procedure and approval by the FDA in 1995. At this point, the surgical procedure was discussed at length in all the popular media, including *The Washington Post*, *The San Francisco Chronicle*, *Newsweek*, and *The New York Magazine*. On 11 October 1999, *Time* magazine published a cover story entitled "The laser fix."

In August 1998, The National Academy of Sciences issued a pamphlet entitled "Preserving the Miracle of Sight: Lasers and Eye Surgery," the stated purpose of which was to show "The Path from Research to Human Benefit." One section describes the first experiments that were done at IBM Research and, subsequently, at Columbia University, leading to the development of PRK [7].

As for the size of the "business" of laser refractive surgery, at a typical cost of $2000/procedure, patients have spent more than $90 billion on PRK and LASIK through the end of 2012.

Srinivasan, Blum, and the author opened the door to this revolution in eye care through their seminal discovery and subsequent transfer of the technology to the medical/surgical profession. The OSA presented this group with the R. W. Wood Prize in 2004 "for the discovery of pulsed ultraviolet laser surgery, wherein laser light cuts and etches biological tissue by photoablation with minimal collateral damage, leading to healing without significant scarring." In 2013, Srinivasan, Blum, and the author received the National Medal of Technology and Innovation from President Obama and the Fritz J. and Dolores H. Russ Prize from the National Academy of Engineering.

References

1. R. Srinivasan and V. Mayne-Banton, "Self-developing photoetching of poly(ethylene terephthalate) films by far-ultraviolet excimer laser radiation," Appl. Phys. Lett. **41**, 576–578 (1982).
2. R. Srinivasan and W. J. Leigh, "Ablative photodecomposition: the action of far-ultraviolet (193 nm) laser radiation on poly(ethylene terephthalate) films," J. Am. Chem. Soc. **104**, 6784–6785 (1982).
3. R. Linsker, R. Srinivasan, J. J. Wynne, and D. R. Alonso, "Far-ultraviolet laser ablation of atherosclerotic lesions," Lasers Surg. Med. **4**, 201–206 (1984).
4. R. Srinivasan, J. J. Wynne, and S. E. Blum, "Far-UV photoetching of organic material," Laser Focus **19**, 62–66 (1983).
5. S. L. Trokel, R. Srinivasan, and B. Braren, "Excimer laser surgery of the cornea," Am. J. Ophthalmol. **96**, 710–715 (1983).
6. F. A. L'Esperance, Jr., "Method for ophthalmological surgery," U.S. patent 4,665,913 (19 May 1987).
7. R. Conlan, "Preserving the miracle of sight: lasers and eye surgery," in Beyond Discovery: The Path from Research to Human Benefit (National Academy of Sciences, 1998). http://www.nasonline.org/publications/beyond-discovery/miracle-of-sight.pdf.

1975–1990

Intraocular Lenses: A More Permanent Alternative

Ian Cox

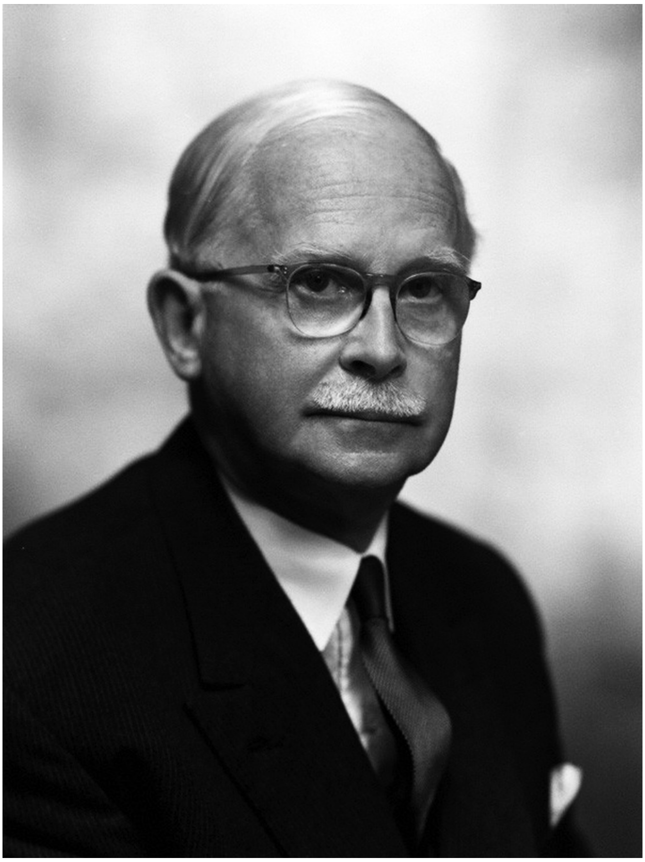
Before the 1950s, cataracts, a loss of transparency of the human lens causing blindness, had been treated using procedures such as "couching" and various forms of intra- and extracapsular lens extraction (ICCE, ECCE). Minimizing surgical complications and attaining good postoperative vision were the primary goals of the surgery. Correction of post-operative aphakia with spectacles was less than satisfactory for patients; their quality of vision was impacted by the magnification, visual aberrations, and field loss inherent in the high-powered positive lenses required to correct the post-surgical eye. Contact lenses provided a superior optical alternative to spectacles, but mobility in the elderly patients typically undergoing cataract surgery was a real problem, as contact lenses needed to be inserted and removed every day.

Sir Harold Ridley (Fig. 1) is universally accepted as the "father" of intraocular lenses (IOL). He was the first to conceptualize a lens that could be surgically implanted in the eye to compensate for the loss of optical power that occurs when the cataractous lens is removed. Noting that fighter pilots injured during the early years of World War II with Plexiglass splinters permanently lodged in their eyes showed no adverse responses, he designed a polymethyl methacrylate (PMMA) optic to replace the cataractous lens in the eye. In 1949 he performed the first surgery to implant a plexiglass intraocular lens. Although the prescription was far from ideal due to errors in the calculation of the refractive index of the natural lens, the surgery was considered a success [1]. Ridley IOLs were used in hundreds of similar surgeries over the next decade, with successful outcomes reported in about 70% of cases. Difficulties in maintaining the lens location in the posterior chamber of the eye and centered on the pupil were the main causes of failure. Amazingly, although a small number of visionary surgeons followed Ridley's lead in the use of intraocular lenses to correct for cataract extraction, it would not be until the late 1980s before it became the preferred method of correction.

From the 1950s through the 1980s, the history of IOL development would be a leap-frogging of technologies in the placement of the IOL in the eye, IOL mechanical design, surgical technique, and diagnostic equipment for measuring the intraocular length of the eye. During this period the lens material of choice was PMMA, with rigid metal or PMMA haptics requiring a large incision size, polypropolene haptics being introduced to help with centering the lens as the capsular bag collapsed during the healing process [2].

In 1984, the first silicone IOL lens, designed by Marzocco and introduced by STAAR, was brought to the marketplace. The huge advantage of this flexible lens was that it could be introduced through the incision into the eye in a folded configuration, allowing a decrease in the surgical incision size. The incision length is related to the induction of post-surgical corneal astigmatism [3], so this signaled the beginning of a drive toward smaller incision sizes that continues to this day. Ridley's original incision was essentially the full diameter of the cornea, while today incisions can be as small as 2 mm, using a dedicated injector to fold and introduce the lens through the incision. It was not until the early 2000s that convergence of these technologies brought a standard of procedure that is the norm in the United States even today [2]. This involves a cataract extraction in the capsule via phacoemulsification under topical intracameral anasthesia. The replacement IOL is a flexible, one-piece lens with a square posterior edge (to reduce posterior capsule opacification), introduced through a 3.0-mm or smaller incision in the cornea and placed fully within the capsular bag, with a slight vault against the posterior surface of the capsule.

Having spent 50 years developing this procedure to be the preferred option for all cataract surgeries, even in children, the industry moved its sights to optimizing the optical performance of IOLs. In 1989 David Atchison identified the considerable increase in spherical aberration created by removing the natural lens and recommended spherical surfaced lens forms that would correct the majority of this aberration [4]. He followed this with the suggestion that using aspheric surfaces would not be beneficial, due to the aberrations induced by tilt and decentration of the final IOL after healing. Not to be deterred, Antonio Guirao and several colleagues, including Pablo Artal and Sverker Norrby, measured the image quality of the normal population with age and then of the typical psuedophakic population. Led by Norrby, an IOL was developed to correct the average spherical aberration of the post-surgical IOL implanted eye. The lens, released to the market by Abbott Medical Optics (AMO) as the TecnisIOL, was designed with an aspheric anterior lens surface and consideration of the typical decentrations that occur with IOL surgical placement and postoperative healing. A rapid response from Alcon provided lenses that corrected a portion of the spherical aberration of the eye and IOL in combination, and Bausch and Lomb provided a spherical aberration-free IOL design, ignoring the spherical aberration inherent in the aphakic eye. All three lenses met with successful use by surgeons around the world, the more technology minded exploring the concept of using all three lenses along with Zernike analysis of corneal topography measurements to determine which lens would come closest to nullifying the spherical aberration of an individual eye.



▲ **Fig. 1.** Sir Harold Ridley, universally accepted as the "father" of IOLs, being the first to devise, produce, and implant the first PMMA IOL. (© National Portrait Gallery, London. Sir (Nicholas) Harold Lloyd Ridley by Bassano Ltd., half-plate film negative, 19 May 1972, NPG x171529.)

The next challenge was correcting near vision in the pseudophakic eye, which of course, has no accommodation after removal of the natural lens. Early attempts at multizonal IOLs for correcting presbyopia demonstrated marginal success due to poor image quality and led to withdrawal from the market by the early 1990s, but in 1997 AMO released a simultaneous refractive multifocal lens (distance, intermediate, and near zones of the design were within the patient's pupil under normal illumination) that gained traction in the marketplace until the early 2000s, when complaints of reduced contrast and halos at night led to a reduction in use [2]. About this time Alcon introduced a diffractive bifocal IOL design, based on patents bought from 3M but updated with a smaller optic zone (only the central 3.6 mm encapsulated the bifocal diffractive element) and an apodized energy profile. The lens had greatest near power at the center of the pupil (equal distance and near), and a shift biased toward distance power moving from the center to the periphery of the optic zone, with all light focused at distance outside the 3.6-mm central diffractive zone. Under its marketed name of ReSTOR, this product met with great enthusiasm when presented to clinicians and continues to grow in popularity, especially in the latest version, which has a lower add power (reduced from +4 D in the original design to +3 D).

AMO responded with a modified refractive multifocal marketed as the ReZoom in 2005, and then released a diffractive design in 2010, which was similar to the Alcon product, without the apodization feature. Although these types of designs are generally successful, some patients do experience reduced contrast, ghosting, and doubling with large pupil sizes, particularly in lenses that are decentered relative to the center of the pupil, as one might expect with designs of this type.

Stuart Cummings, a surgeon, observed in 1989 that patients who had plate haptic silicone IOLs inserted often showed better near reading performance than those fitted with other conventional loop haptic IOL designs, leading him to invent a lens specifically designed to optimize this feature. By adding a weakened portion or "hinge" to the plate haptic, the silicone lens was designed to bend under the intraocular forces occurring with ciliary muscle contraction during accommodation. In this way, the optics of the lens were traditional monofocal spherical surfaces, but good image quality could be provided at both distance and near as the optic of the lens moved forward with the accommodative response. Brought to the market under the tradename Crystalens in 2005, this lens was the first, and is still the only, IOL to have the claim approved by the FDA that it demonstrates "accommodation" of up to 1 D. The exact mechanism of action has not been verified, but it is most probably a combination of optic displacement, optic tilt, and optic zone distortion brought about by the accommodative forces of the eye increasing the depth of field. Regardless of the mechanism, clinical studies have shown superior near vision over monofocal lenses, while maintaining equivalent distance visual acuity.

Correction of postoperative astigmatism induced by surgery was always an issue with cataract surgery, as large incisions closed by sutures led to significant changes in corneal topography [3]. Typically these changes would be corrected by progressive spectacles worn by the pseudophakic patient postoperatively. However, the acceptance of multifocal IOLs through the 2000s in conjunction with small, sutureless incision sizes led to an expectation from many patients that they could spend most of their waking hours without a distance spectacle correction. This paradigm opened the demand for toric IOLs in those patients who had significant corneal astigmatism prior to cataract surgery. Although offered to the industry in 1994 by STAAR on their plate silicone lens platform, significant adoption of toric IOLs only began with the introduction of the Acrysof Toric IOL by Alcon in 2005. Although optically the design is straightforward, a successful toric IOL must demonstrate stability of the cylinder axis from lens placement at the time of surgery until complete healing 3 to 6 months postoperatively. This lens, along with competitor offerings, typically shows stability that makes the use of toric lenses a benefit in eyes with 1.25 D of astigmatism or greater postoperatively.

IOLs have come a long way since their beginnings in 1949, and today they are the preferred method of correction following cataract surgery regardless of patient age or refractive status.

References

1. H. Ridley, "Intra-ocular acrylic lenses—past, present and future," Trans. Ophthalmol. Soc. UK **84**(5), 5–14 (1964).
2. J. A. Davison, G. Kleinmann, and D. J. Apple, "Intraocular lenses," Chap. 11 in *Duane's Ophthalmology* (CD-ROM), W. Tasman and E. A. Jaeger, eds. (Lippincott Williams & Wilkins, 2006).
3. K. Hayashi, H. Hayashi, F. Nakao, and F. Hayashi, "The correlation between incision size and corneal shape changes in sutureless cataract surgery," Ophthalmology **102**, 550–556 (1995).
4. D. A. Atchison, "Optical design of intraocular lenses. I. On-axis performance," Optom. Vis. Sci. **66**, 492–506 (1995).

Spectacles: Past, Present, and Future

William Charman

Spectacles probably have a longer history than any other optical device, apart from magnifiers, and their development has continued throughout the era of The Optical Society (OSA). A fascinating aspect of this history is that spectacle lens design and technology involve not only optical solutions to the visual needs of the wearer but also considerations of comfort, fashion, and appearance. In particular, the diameter of lens required to fit any frame may put serious constraints on the optical characteristics of the lens.

The optics of the human eye should form an image of the outside world on the light-sensitive retina. Since objects of interest may lie anywhere between distant and relatively close distances of the order of arm's length or less, either the depth of focus of the eye must be very large or, more realistically in view of the eye's relatively large maximal numerical aperture, ~0.25, an active focusing mechanism is required. Focusing is achieved by active changes in the shape of the elastic crystalline lens, a process known as accommodation. With accommodation relaxed, the eye ought to be focused for distance, when it is called emmetropic.

Unfortunately, our evolutionary development has left us with two problems. First, the ocular dioptics may not form a sharply focused image of distant objects, so that the eye suffers from ametropia. If the optics are too powerful, the image lies in front of the retina, and the eye is myopic ("short-sighted"); if too weak, the image lies behind the retina and the eye is hyperopic (often erroneously called "long-sighted"). Evidently the myopic eye can focus clearly on near objects and the hyperopic eye may be able to increase its power by accommodation to focus both distant and some near objects. The second problem is that while accommodation was adequate to the needs of our short-lived ancestors, most of us are now living too long for accommodation to remain effective in the later part of life. The objective amplitude of accommodation (i.e., the maximum change in ocular power) for each of us declines steadily from the early teenage years to reach zero at about 50, when the individual becomes fully presbyopic. Thus, older uncorrected emmetropes and hyperopes inevitably have poor near vision, although myopes have less difficulty. Almost all older individuals need some form of optical assistance if they are to see both distant and near objects clearly, the only exceptions being a few happy anisometropic individuals, having one near-emmetropic eye and one mildly myopic eye.

By 1916, at the time when the OSA was founded, basic spectacle lens design was reasonably well understood. A variety of types of bifocals were available, including the fused form, where the bifocal near segment was made of flint glass and the distance carrier was made of crown so that the "add" effect could be obtained with a lens having no surface discontinuities. Prisms had been introduced by Von Graefe and Donders to help those with convergence problems. Tints of various colors and transmittances were available (indeed, as early as Christmas Eve 1666, the great diarist Samuel Pepys was writing "I did buy me a pair of green spectacles, to see whether they will help my eyes or no"). After seven centuries of development, could spectacle lenses be improved further?

Spectacle lens design and the materials used have, in fact, advanced to a surprising degree during the "OSA century." The earliest relevant paper in the OSA's brave new flagship publication, *Journal of The Optical Society of America*, appeared in the first volume under the title "The reflected images in spectacle lenses" [1]. These reflections may interfere with the wearer's vision but are generally considered to be most important from the cosmetic point of view. Since for normal incidence the reflectance at the surface of a lens of refractive index n is $(n-1)^2/(n+1)^2$, the problem increases as the lens index is raised. Single-layer and multi-layer coatings have, in recent decades, provided a solution, but questions remain on the optimal coating characteristics, since under conditions of spectacle use fingerprints and other dirt may, on the lens, be more obvious on the coated lens, and regular cleaning is required. It is, incidentally, of interest that as late as 1938 Tillyer, in a discussion on optical glasses given at an OSA symposium on optical materials, still thought it worth commenting "more light gets through the lens when it is tarnished slightly"—an earlier, less controlled form of lens coating!

The question of lens index is also, of course, of great importance in relation to lens thickness and the consequent appearance of the spectacles when worn. Surface power is given by $(n-1)/r$, where r is the surface radius. Thus, for any required corrective power, the difference between the two surface curvatures of a meniscus spectacle lens will be reduced if its index is increased. This means that a positive lens can have smaller central thickness and a negative lens will have reduced edge thickness for any given lens diameter. This is of particularly cosmetic value for high myopes wanting a frame that demands a large lens diameter. Depending upon the material density, the weight of the thinner lens may also be reduced. Thus, over recent decades there have been continuing and successful attempts to produce materials of higher refractive index, in both glass and plastic. Whereas traditional crown and flint glasses had indices of 1.52 and 1.62, respectively, materials are now available with indices up to 1.9.

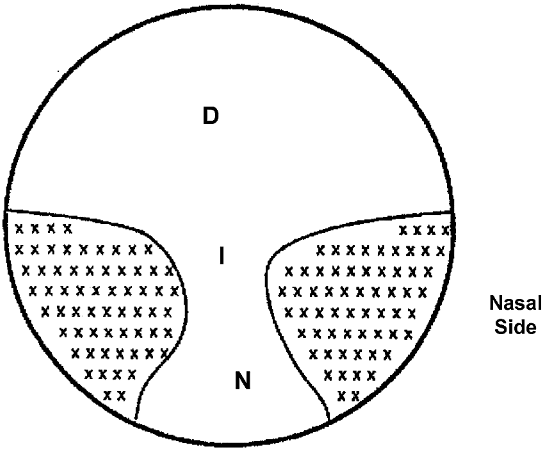
Refractive index and density are, however, not the only consideration with lens materials. Dispersive characteristics are also important, since when directing the visual axis away from the lens center the wearer is effectively looking through a prism, resulting in transverse chromatic aberration and color fringing around objects. Thus, as well as having high index and low density, the ideal lens material should have as high a constringence (Abbe number, V-Value) as possible. Currently glasses of refractive index 1.8 have a constringence of about 35.

A major advance in materials was the appearance of plastic lenses. Although polyethyl methacrylate (PMMA, Plexiglass, Perspex) had been introduced before the second world war, it was relatively soft and easily scratched. The breakthrough came with a wartime development, CR39, a polymerizable, thermosetting plastic with a refractive index (1.498) similar to that of crown glass and a V-value of 58. Importantly, it had better scratch resistance than PMMA, a high impact resistance, and half the density of crown glass. The first ophthalmic lenses in the material were produced by Armorlite in 1947. Lenses can be either surfaced or molded. Demands for still higher impact resistance led to the introduction of polycarbonate lenses in the late 1950s, first for safety eyeware and later, as optical quality improved, for all powers of ophthalmic lens. Polycarbonate is a thermoplastic, and lenses can again be made by either molding or surfacing techniques. Its index (1.586) is a little higher than crown glass but its V-value (30) is lower: since the scratch resistance is not high, the surface is usually protected by a hard coating, such as thermally cured polysiloxane. The specific gravity and UV transmittance are low. Other higher index plastics are now available. Various hard and anti-reflection coatings can be applied to all these plastic lenses, whose many attractive features have given them a dominant position in the spectacle market. Ultimately gradient-index media may find a role in spectacle lens design [2].

From the design point of view, the advent of computers has allowed the impact of aspherization on the performance of single-vision lenses to be explored in considerable detail [3]. Such work has revealed that aspherization widens the range of lens forms that yield zero oblique astigmatism as compared to those lying on the Tscherning ellipse. Modern ray-tracing techniques have also greatly benefited the design of progressive addition (varifocal) lenses. These are lenses for presbyopes in which the discrete power zones of traditional bifocals and trifocals are replaced by a smooth variation in power across the lens surface, from that appropriate for distance vision to that for near, with good vision for intermediate distances between the distance and near zones and an absence of visible dividing lines on the lens surface. First proposed by Aves in 1907, with his "elephant's trunk" design, the first successful lenses of this type were the French Varilux designed by Maitenaz (Essilor) and, in the U.S., the Omnifocal (Univis). Since then numerous variations have been produced. Optically, the challenge is that the shorter the progressive corridor between stable distance and near corrections, the narrower the corridor and the greater the unwanted astigmatism in neighboring lens areas (Fig. 1). Since the visual axes converge during near vision, separate right and left eye lenses are required. Moreover the "ideal" lens depends on such factors as the extent to which the individual patient moves the eyes or the head when changing fixation. Thus, the concept of "customized" lenses has been introduced, where details of the design depend upon the characteristics of the individual wearer and the frame used. The manufacture of such lenses is only possible through the recent availability of digital surfacing or "freeform" technology. An obvious downside is that the advantages of customization may be destroyed if the lenses are in the incorrect position as a result of frame movement or distortion.

While neutral and color-tinted lenses have been available for many centuries, with progressive refinement in bulk, coated, or laminated forms, one striking innovation in the OSA era was the introduction, by Corning in the mid-1960s, of photochromic lenses. These actively change their transmittance in response to the ambient light level, obviating additional prescription sunglasses. The original glass-based photochromics relied on silver halide, in which electron exchange under the influence of high levels of short-wavelength light yielded opaque colloidal metallic silver. The resultant loss in transmittance was reversed when the light levels lowered, with transition times of the order of a few minutes. Subsequent advances have resulted in more stable lenses with shorter transition times and photochromic plastics using organic dyes.

One specialized area of spectacle use is for low-vision patients, who require magnification for either distance or near tasks. Ellerbrock [5] gave a valuable account of the aids available at that time, and the OSA later honored an outstanding practitioner in the field, Louise Sloan, by the award of its Tillyer Medal in 1971 [6] (Fig. 2). The question of whether wearers of bioptic spectacles, with their limitations on field of view, should be allowed to drive remains controversial. "Press-on" plastic Fresnel lenses and prisms have found application in patients with binocular vision problems such as squint.



▲ **Fig. 1.** Zones of a progressive addition lens (PAL). The distance D and near N zones are connected by a progressive intermediate zone (I). Areas of poor vision because of unwanted surface astigmatism are shown by shading. (Reproduced with permission of [4]. Copyright 1993, The Optical Society.)

▲ **Fig. 2.** Louise Sloan receiving the Tillyer Medal in 1971.

What does the future hold? One challenge is the search for a full-aperture lens of variable power for the correction of presbyopia, so that the accommodational ability of the young eye can be mimicked. While multi-lens "zoom" spectacles exist, their appearance makes them unacceptable to all except a minority of presbyopes. Variable-power lenses with a fluid reservoir enclosed by a flexible membrane, so that the surface curvature can be varied by pumping liquid in or out, have a long history but have so far found only a limited market. Alvarez lenses, consisting of two closely spaced component lenses with surfaces following a cubic equation that are translated laterally with respect to each other, have found some application recently. Like membrane lenses they are difficult to incorporate into standard frames. Possibly more promising are electrically switched devices, such as liquid-crystal refractive or diffractive lenses, but the latter suffer from the problem of large amounts of transverse chromatic aberration. The search continues.

Finally, there is continuing interest in the interaction of spectacles with the growth of the eye and the development of refractive error. In recent decades the prevalence of myopia has increased, particularly in many Asian countries, presumably associated with lifestyle changes for those involving near work or outdoor activity. Can a child's wearing of suitable spectacles eliminate, or at least reduce, these myopic changes? Animal experiments suggest that the axial length of the growing eye is affected by lens wear and that peripheral as well as axial imagery are of importance. Thus current studies are exploring the possible beneficial effects of bifocal or other lenses to relieve accommodation demand and lenses that modify the pattern of peripheral refraction.

Many spectacle challenges remain for future members of the OSA!

References

1. W. B. Rayton, "The reflected images in spectacle lenses," J. Opt. Soc. Am. **1**(5–6), 137–148 (1917).
2. S. P. Wu, E. Nihei, and Y. Koike, "Large radial graded-index polymer," Appl. Opt. **35**, 28–32 (1996).
3. D. A. Atchison, "Spectacle lens design: a review," Appl. Opt. **31**, 3579–3585 (1992).
4. C. W. Fowler, "Method for the design and simulation of progressive addition spectacle lenses," Appl. Opt. **32**, 4144–4146 (1993).
5. V. J. Ellerbrock, "Report on survey of optical aids for subnormal vision," J. Opt. Soc. Am. **36**, 679–695 (1946).
6. L. L. Sloan, "Optical magnification for subnormal vision: historical survey," J. Opt. Soc. Am. **62**, 162–168 (1972).

Major Milestones in Liquid Crystal Display Development

Shin-Tson Wu

The earliest display of moving images was the motion picture projector, in which light from a bright lamp was passed through an image on a film that was then imaged onto a screen. In the 1920s and 1930s the first black and white television broadcasts were made and viewed on small black-and-white cathode ray tube displays. Such a display was achieved by writing a visible image on a phosphor screen with an electron beam. It required a vacuum tube and high voltage electronics, yet it produced a reasonable image. Over time cathode ray tube displays became larger and capable of color images. They also became very heavy, bulky, and power hungry, though they had good color rendition. However, they were all there were, and the industry developed color CRTs with screen sizes as large as 1 m in diagonal dimension. Alternative displays were tried such as plasma screens (an array of tiny, energy-hungry plasmas that excited special phosphors for each color that quickly were bleached by the UV in the plasma) or micro-mirror scanner displays. However, all of these were supplanted by the advent of the liquid crystal display, the LCD. Today these displays dominate the display marketplace due to their ability to be used in all sizes, from as small as a wristwatch to over 2.8-m-diagonal television screens. LCDs can be reflective, requiring just ambient light to be viewed, transmissive, requiring a backlight to enable viewing, or transreflective, in which a pixel is split into reflective and transmissive subpixels. In either case their advantages of light weight, lower energy demand, and scalability have won LCDs a dominant place in today's display marketplace. This essay explores how that happened.

Liquid crystal is a mesogenic phase existing between crystalline solid and isotropic liquid. In 1888, Austrian botanist Friedrich Reinitzer and German physicist Otto Lehmann discovered such an anisotropic liquid crystal. However, in the early days only a few compounds with a liquid crystal phase were available, and their melting points were quite high. Moreover, to utilize its large optical anisotropy the liquid crystal has to be aligned and an external field applied. Before the optically transparent and electrically conductive indium-tin-oxide (ITO) film was available, an alternative way to align a liquid crystal was by applying a magnetic field. Therefore, in the first few decades major research focused on magnetic-field-induced molecular reorientation effects. But the electromagnet required to align the liquid crystals was too bulky to be practically useful. Then in the 1930s Russian scientist V. Fréedericksz and colleagues started to investigate the electro-optic effects in nematic liquid crystals. Some basic concepts were formulated such as the Fréedericksz transition threshold and order parameter, which described the crystalline state of a liquid crystal. In the 1950s and 1960s, the dynamic behavior of a liquid crystal cell subjected to an external force, such as a magnetic field or electric field, was investigated by C. W. Oseen, F. C. Frank, J. L. Ericksen, and F. M. Leslie. These concepts and models provided the foundation for the rapid development of the useful electro-optic devices that followed.

In the 1960s, American scientists George Heilmeier, Richard Williams, and their colleagues at RCA (Radio Corporation of America) Labs developed the dynamic scattering mode and demonstrated the first LCD panel [1]. This opened a new era for electronic displays. Heilmeier was credited with the invention of the LCD. In 2006, he received the OSA Edwin H. Land Medal, and in 2009 he was inducted into the National Inventors Hall of

Fame. However, the dynamic scattering LCD, which utilized the electric-current-induced electrohydrodynamic effect, was intrinsically unstable. Also, its contrast ratio was poor and power consumption was high. As a result, it had a short life and was ultimately abandoned as a practical display technology.

In the 1970s, to overcome the instability, poor contrast ratio, and high operation voltage of the dynamic scattering mode display, Martin Schadt and Wolfgang Helfrich, and James Fergason independently, invented the twisted nematic (TN) effect and steered LCD in a new and productive direction. TN is regarded as a major invention of the twentieth century. In 1998, James Fergason was inducted into the National Inventors Hall of Fame. In 2008 Schadt, Helfrich, and Fergason received the IEEE Jun-Ichi Nishizawa medal in recognition of their outstanding contribution.

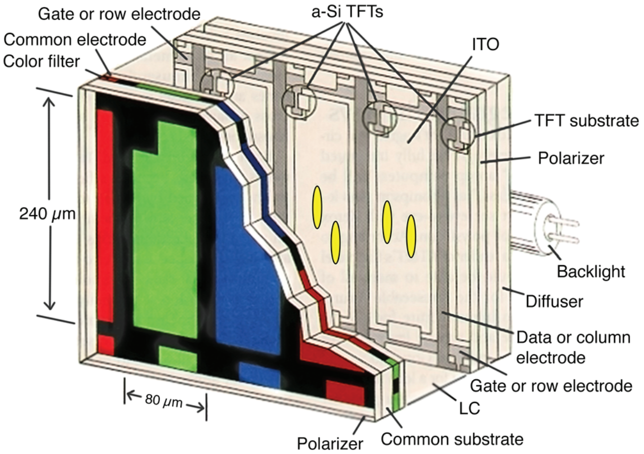
Also in the 1970s, a landmark equally important to TN was the development of stable liquid crystals called cyanobiphenyls by George Gray's group at Hull University [2]. Amazingly, these positive dielectric anisotropy ($\Delta\varepsilon \sim 15$) materials are still being used in some wristwatches and calculators in 2016. Meanwhile, to obtain a uniform domain new liquid crystal alignment techniques were developed. Among them, buffed polyimide deserves special mention because it enables large panel LCDs to be fabricated. This technique is still commonly used in modern LCD fabrication lines. Liquid crystals need a small pre-tilt angle (3°–5°) to guide their reorientation direction when activated by an electric field. Otherwise, different domains could be formed, which caused spatially inhomogeneous electro-optic behaviors. In addition to TN, vertical alignment (VA) and in-plane switching (IPS) were invented in the 1970s. In TN and VA cells, the electric field is in the longitudinal direction, while in an IPS cell the electric field is in the lateral direction, also called the fringing field. These three modes form the bases of modern LCD technologies. TN is used in notebook computers and personal TVs in some aircraft because of its low cost and high transmittance; multi-domain VA is widely used in high-definition TVs because of its unprecedented contrast ratio; and IPS is commonly used in mobile displays, such as iPhones and iPads, because of its robustness to external mechanical pressure allowing use in touch screens.

Another crucial development in the 1970s was the thin film transistor liquid crystal display (TFT LCD) led by Bernard Lechner at RCA and Peter Brody's group at Westinghouse. In 1972, a group at Westinghouse led by A. G. Fisher demonstrated that a color TV could be made by integrating red (R), green (G), and blue (B) spatial color filters with liquid crystal pixels as intensity modulators [3]. Each color pixel was independently controlled by a TFT. This combination of TFT and LCD enabled high information content and became the foundation of today's display industry. In 2011, three TFT pioneers—Bernard Lechner, Peter Brody, and Fang-Chen Luo—received the IEEE Jun-Ichi Nishizawa medal, and in 2012 Heilmeier, Helfrich, Schadt, and (the late) Brody received the prestigious National Academy of Engineering's Charles Stark Draper Prize to recognize their engineering development of LCD utilized in billions of consumer and professional devices.

The early TFTs developed by Brody and his colleagues were based on cadmium selenide (CdSe), which was never commercialized because of high off-current and reliability issues. Today, most LCDs use silicon TFTs: amorphous silicon for large panels [>10-in. (25 cm) diagonal], poly-silicon for small-to-medium panels such as iPhones/iPads, and single-crystal silicon for micro-displays. Recently, oxide semiconductors, e.g., $InGaZnO_2$ with mobility about 20× higher than that of amorphous silicon, have been attempted in TFT LCDs by major display producers. The high mobility of oxide semiconductors helps to shrink TFT feature size, which in turn leads to a larger aperture for higher backlight throughput.

In the 1980s, passive matrix and active matrix addressed LCDs were pursued in parallel. In the passive matrix camp, a new LC mode called super-twisted nematic (STN; twist angle >90°) was developed to steepen the voltage-dependent transmittance curve to increase information content. However, the viewing angle, contrast ratio, and response time of STN are far from satisfactory. In the active matrix camp, Seiko, Epson, and several Japanese display leaders invested heavily in active matrix TFT-LCD production facilities. In the meantime, new high-resistivity fluorinated liquid crystals were developed; this technology is required for active matrix operation to avoid image flickering. After nearly a decade of fierce competition, active matrix outperformed passive matrix and is commonly used in display products.

Figure 1 shows the device structure (one color pixel consisting of three RGB sub-pixels) of a TFT-LCD. LCD is a non-emissive display, so it requires a backlight or edge light, such as a cold cathode fluorescent lamp (CCFL) or a light-emitting-diode (LED) array. A thin liquid crystal layer is sandwiched between the active matrix substrate and color filter substrate, functioning as a spatial light modulator. Each sub-pixel is controlled by a TFT switch.



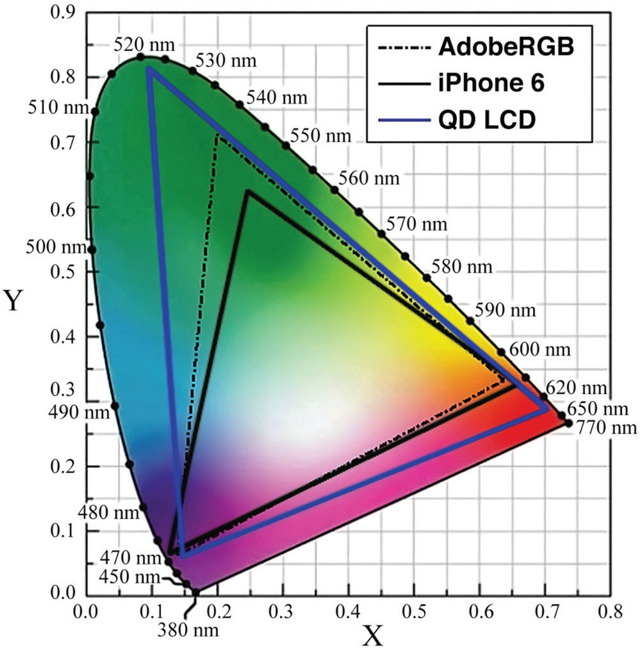
▲ Fig. 1. Device structure of a color pixel of thin-film-transistor LCD.

An important advancement in the 1990s was wide-view technology. Liquid crystal is a birefringent material, so its electro-optic property depends on the viewing direction. This problem gets worse as the panel size increases. To widen the viewing angle, two major approaches were undertaken: (1) multi-domain structure, e.g., four domains, and (2) phase-compensation films to reduce light leakage at oblique angles. To create four domains, zigzag electrode patterns were used. The viewer sees the average effect from four domains with size around 100 μm. Therefore, the viewing angle is widened dramatically. Once the viewing angle issue was overcome, there was a huge movement toward producing large-panel LCDs by Korean and Taiwanese manufacturers, in addition to those in Japan.

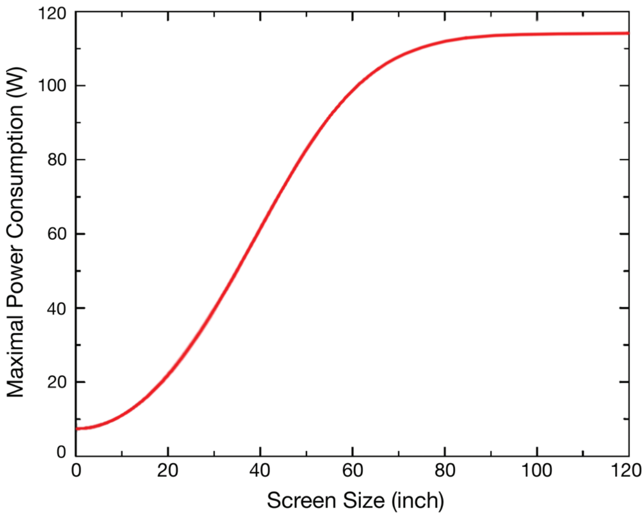
In the 2000s, in addition to large screen sizes and high resolution, LCD received two important enhancements: LED backlights and touch panels. The traditional backlight was a CCFL. It has a narrow green emission, but the red and blue are broad. As a result, some blue–green and yellow–red emissions leak through the corresponding blue and red color filters, so that the color gamut is limited to ~75%, similar to that of a CRT. To improve color saturation and reduce power consumption, two types of LED backlight were considered: white LEDs and RGB LEDs. White light can be generated by using a blue LED to excite yellow-emitting phosphors or combining RGB LEDs. The former approach is quite efficient, but its yellow emission is quite broad. Consequently, the color gamut is also limited to ~75%. The RGB approach greatly extends the color gamut to over 120%; however, it requires three driving circuits for the RGB LEDs. Moreover, there is a so-called “green gap” in the LED industry. That means there is limited choice for green LEDs in terms of color and efficacy. Both approaches were utilized by some major LCD developers, but eventually white LEDs won out. Nowadays, benefiting from progress in the general lighting industry, the efficacy of white LEDs has exceeded 100 lm/W. The touch-panel LCD was another important technological development in the 2000s. The Apple iPhone and iPad are examples of touch-panel LCDs. Numerous touch technologies were developed, including resistive, capacitive, surface acoustic wave, infrared, and optical.

In 2004, as a consequence of the rapid growth in the display industry, IEEE and OSA jointly launched a new journal, called the *Journal of Display Technology* (JDT). The author served as the founding editor-in-chief. The scope of JDT covers all aspects of display technologies, from understanding the basic science and engineering of devices, to device fabrication, system design, applications, and human factors.

In the 2010s, major research and development focused on faster response time, more vivid colors, higher resolution, larger panel sizes, curved displays, and lower power consumption. CRT is an impulse-type display; once the high-energy electrons bombard phosphors, the emitted light decays rapidly. Therefore, the displayed images do not remain at the viewer’s eye, which means the images are clear. The only problem is that the frame rate should be fast enough (~120 Hz) to minimize image flickering. Unlike CRT, TFT-LCD is a holding-type display. Once the gate channel is open, the incoming data signals charge the capacitor and stay there until the next frame comes. Therefore, TFT-LCD is ideal for displaying static images, such as paintings. When displaying fast-moving objects, the



▲ **Fig. 2.** Simulated color gamut of the iPhone 6 and quantum-dot-enhanced LCD.



▲ **Fig. 3.** Maximum power consumption set by Energy Star 6. Aspect ratio: 16:9.

holding-type TFT LCD causes image blurs. To suppress image blurring, we can increase the frame rate, blink the backlight to make CRT-like impulses, and reduce the LC response time, which is governed by the visco-elastic coefficient of the LC material and the square of the cell gap. With continued improvement in developing low-viscosity LC materials and advanced manufacturing technology to control the cell gap at ~3 μm, the response time can be as small as ~4 ms.

Another issue for LCDs to overcome is color. Most LCDs use single-chip white LED backlighting: a blue InGaN LED to pump a yellow phosphor (cerium-doped yttrium aluminum garnet: Ce:YAG). This approach is efficient and cost effective, but its color gamut is ~75% and cannot faithfully reproduce the natural colors. Recently, quantum-dot (QD) LEDs are emerging as a new backlight source. Resulting from the quantum confinement effect, a QD LED exhibits high quantum efficiency, narrow-emission linewidth (~30 nm), and controllable emission peak wavelength. In comparison with conventional backlight solutions, QD backlight offers a wider color gamut. Figure 2 shows the simulated color gamut of the iPhone 6 (which uses white LED) and QD-enhanced LCD, whose color gamut is over 115% NTSC in CIE 1931 color space [4].

Power consumption affects the battery life of a mobile display and the electricity bill of a LCD TV. To be eco-friendly, Energy Star 6 sets the maximum power consumption for a given display size regardless of which technology is used. Figure 3 shows the maximum power consumption of a display panel with 16:9 aspect ratio. For example, the maximum power consumption of a 60-in. (1.52 m) diagonal HDTV (resolution 1920 × 1080) is ~100 W. As the resolution density keeps increasing, the TFT aperture ratio is reduced and power consumption is increased. To reduce power consumption, several approaches can be considered, such as a more efficient LED backlight, backlight recycling, a high-mobility oxide semiconductor to increase the TFT aperture ratio, and color sequential display to remove spatial color filters.

In the past five decades, we have witnessed the amazing progress of liquid crystal displays from concept proof to widespread applications. The technology trend is to go with a thinner profile, flexibility and bendability, lighter weight, more vivid color, lower power consumption, and lower cost.

References

1. M. Schadt, "Milestone in the history of field-effect liquid crystal displays and materials," Jpn. J. Appl. Phys. **48**, 03B001 (2009).
2. J. W. Goodby, "The nanoscale engineering of nematic liquid crystals for displays," Liq. Cryst. **38**, 1363–1387 (2011).
3. Y. Ishii, "The world of liquid-crystal display TVs—past, present, and future," J. Disp. Technol. **7**, 351–360 (2007).
4. Z. Luo, Y. Chen, and S. T. Wu, "Wide color gamut LCD with a quantum dot backlight," Opt. Express **21**, 26269–26284 (2013).