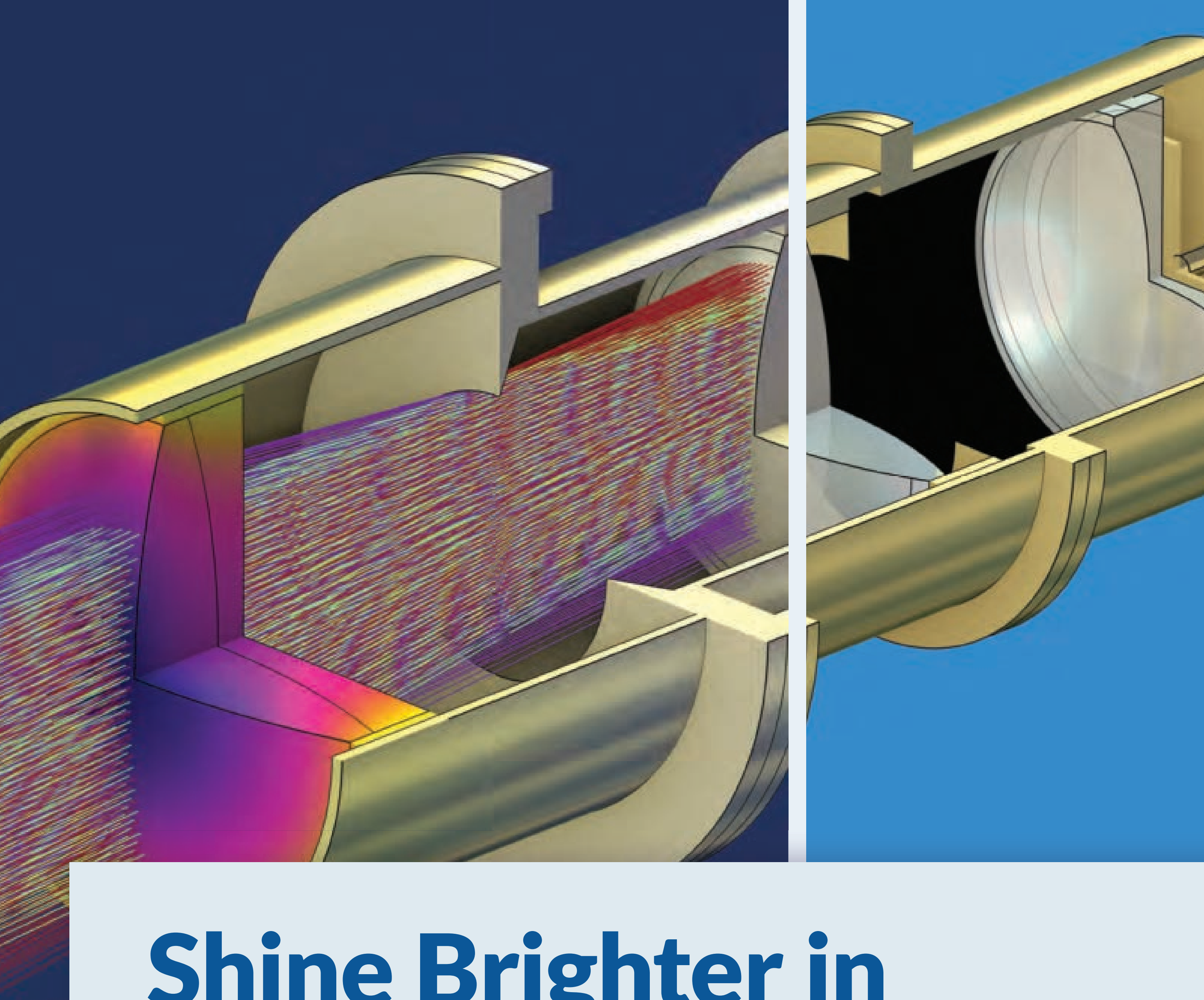


PHYSICS TODAY

August 2023 • volume 76, number 8

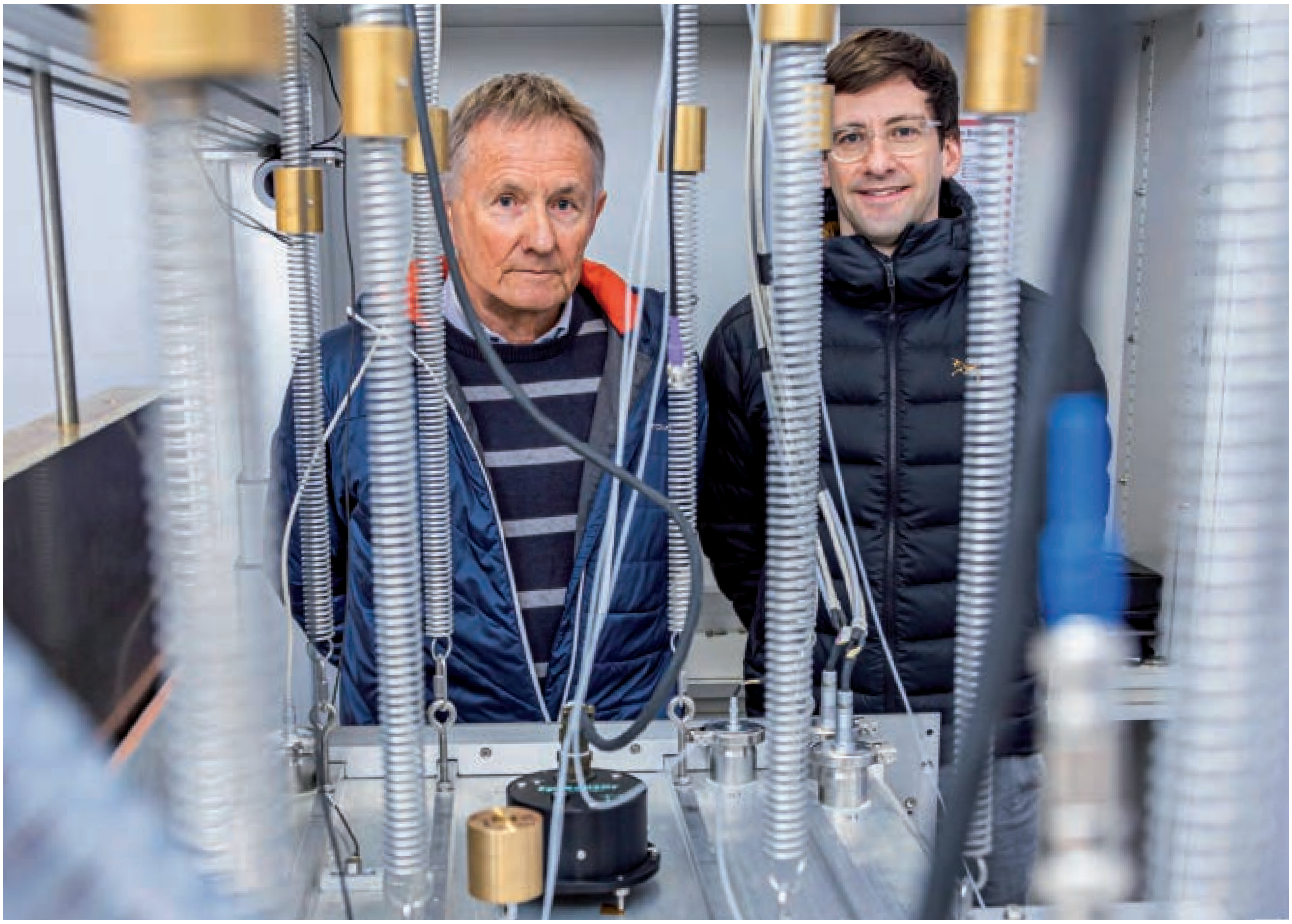
A publication of the American Institute of Physics

QUANTUM TECHNOLOGIES

Advances in solar telescopes

ITER seeks to overcome latest setbacks

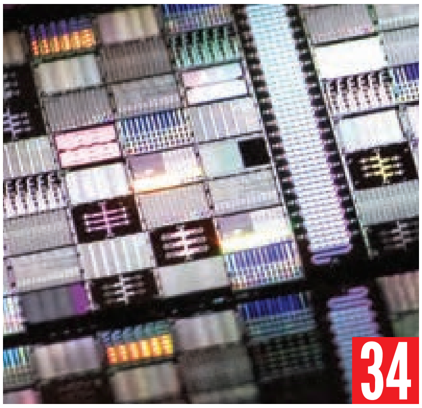
Tapping tidal energy



Prof. Jürg Dual and Dr. Tobias Brack, ETH Zürich

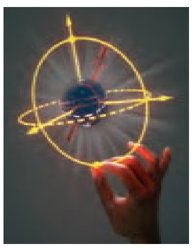
PHYSICS TODAY

August 2023 | volume 76 number 8



FEATURES

SPECIAL FOCUS ON QUANTUM TECHNOLOGIES



ON THE COVER: As quantum computation and other quantum applications speed toward real-world uses, understanding the underlying physics is important for overcoming the technologies' performance challenges. On **page 26**, Christopher Anderson and David Awschalom explore progress on and opportunities for quantum systems based on solid-state defects. And on **page 34**, José Aumentado, Gianluigi Catelani, and Kyle Serniak discuss one of the primary limits on superconducting qubit performance. (Image courtesy of Peter Allen.)

Recently on PHYSICS TODAY ONLINE

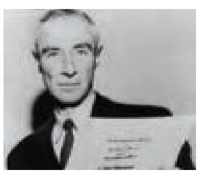
www.physicstoday.org

MAST/STSCI/JWST/T. RECTOR

Astro image processing

It's as much an art as it is a science to transform the raw, black-and-white IR filter data from the *James Webb Space Telescope* into spectacular images. PHYSICS TODAY's Jennifer Sieben investigates how astronomers do it and applies those lessons to create her own full-color images.

physicstoday.org/Aug2023a

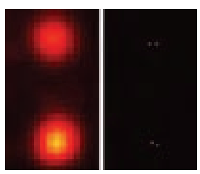


AIP ESVA

Oppenheimer

Now the subject of a major Hollywood movie, J. Robert Oppenheimer has appeared frequently in PHYSICS TODAY. We highlight past coverage of the late Los Alamos director, from the image of his hat on the cover of our first issue in 1948 to more recent news about his US government security clearance.

physicstoday.org/Aug2023b



LLNL

Microscopy advance

Superresolution microscopy enables optical microscopes to resolve objects smaller than the diffraction limit. Using a new technique, researchers can achieve subnanometer resolution by tagging target molecules with several distinct fluorescent labels and averaging the results of multiple imaging sequences.

physicstoday.org/Aug2023c

PHYSICS TODAY (ISSN 0031-9228, coden PHTOAD) volume 76, number 8. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Periodicals postage paid at Huntington Station, NY, and at additional mailing offices. POSTMASTER: Send address changes to PHYSICS TODAY, American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Views expressed in PHYSICS TODAY and on its website are those of the authors and not necessarily those of AIP or any of its member societies.

PHYSICS TODAY

www.physicstoday.org

DEPARTMENTS

14

18

48

SUBSCRIPTION QUESTIONS? +1 800 344-6902 | +1 516 576-2270 | **ptsubs@aip.org**

Address
American Center for Physics
One Physics Ellipse
College Park, MD 20740-3842
+1 301 209-3100
pteditors@aip.org

PhysicsToday **@physicstoday**

AIP

Member societies
ACA: The Structural Science Society
Acoustical Society of America
American Association of Physicists in Medicine
American Association of Physics Teachers
American Astronomical Society
American Meteorological Society
American Physical Society
AVS: Science & Technology of Materials, Interfaces, and Processing
Optica (formerly The Optical Society)
The Society of Rheology

Other member organizations
Sigma Pi Sigma Physics and Astronomy Honor Society
Society of Physics Students

SPECIAL OFFER

Post Your REU at SPS Jobs, FREE

SPS Jobs offers FREE postings to employers recruiting seasonal REU's and interns.

Positions posted on SPS Jobs will also appear on:

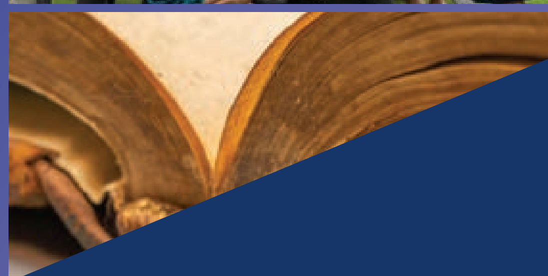
- Physics Today Jobs
- American Association of Physics Teachers
- AVS Science and Technology

Don't Miss Out!
Students are on the hunt for summer opportunities now.

Get started now!

1. Create or login to your SPS Jobs account
   jobs.spsnational.org/employers
2. Select "Summer Research/Internship" Job Level to access the no-cost posting offer
3. Post Your REU

PLEASE NOTE: Valid intern-level opportunities are defined as limited-term (up to 12 weeks) employment for current undergraduates or recent Bachelor-degree recipients (within one year), with compensation in the form of a modest salary/stipend or academic credits.
Questions/Validation - contact spsjobs@aip.org



Physicists need to be talking about nuclear weapons

At the beginning of the nuclear era, many of the leading physicists of the 20th century—including Albert Einstein, Niels Bohr, Enrico Fermi, Werner Heisenberg, Hans Bethe, and J. Robert Oppenheimer—and many of the younger physicists involved in the US's secret nuclear weapons project during World War II discussed the consequences of their work and issued warnings about the dangers of a nuclear arms race with the Soviet Union. Currently, however, despite Russia's threats of nuclear use and a developing nuclear arms race with China, the US physics community is having little such discussion.

That is, perhaps, because until recently and in contrast with the Cold War nuclear arms race, nuclear weapons work in the US has been focused primarily on maintaining and refurbishing the existing nuclear warhead stockpile. But the consequences of using a significant fraction of the global stock of nuclear weapons remain existential—on a par with the long-term consequences of climate change, which is attracting much more attention and debate. It is time for the physics community to renew its engagement with the nuclear weapons policy debate.

Today about 10 000 nuclear warheads are in military service in nine nations (and about 3000 more are waiting to be dismantled).[1] They have an average explosive energy an order of magnitude larger than that of the warhead that produced over 100 000 civilian deaths in Hiroshima, Japan (which had an energy equivalent to about 15 000 tons of chemical explosive).[2] Nearly 2000 warheads are on alert status, ready to be launched within minutes of an order being received.[3] They could kill hundreds of millions of people directly. Indirect effects, such as global famine from crop failures from a multiyear climate cooling due to a stratospheric smoke pall produced by nuclear-caused firestorms, could kill billions.[4]

It is difficult to grasp that humans have allowed themselves to be subject 24/7 to a potentially civilization-ending threat that a single political leader could trigger. Contrary to their commitment in the Treaty on

TIM MENZIES/CC BY-SA 2.0

MODEL OF MIRVS—multiple independently targeted reentry vehicles—around the third stage of a US Trident submarine-launched ballistic missile. Each cone-shaped reentry vehicle houses a nuclear warhead with a yield of up to 455 kt. The diameter of the Trident missile is 2.11 m.

the Non-Proliferation of Nuclear Weapons "to pursue negotiations in good faith on effective measures relating to cessation of the nuclear arms race at an early date and to nuclear disarmament," the five permanent, veto-wielding members of the United Nations Security Council—the US, Russia, the UK, France, and China—all of which are party to the treaty, are upgrading their nuclear arsenals.

Physicists are essential to the nuclear weapons complexes: They perform R&D to maintain and improve nuclear warheads and their delivery systems. Physicists are employed in the multibillion-dollar US Stockpile Stewardship Program (SSP), which aims to ensure "safety, security, reliability, and effectiveness" of the nuclear weapons stockpile.

Minimizing the possibilities of nuclear accident (safety) or unauthorized use (security) is essential. The SSP's large focus on nuclear weapons' reliability and effectiveness deflects attention, however, from the fact that the weapons are unusable by responsible leaders. It is difficult to imagine that a rational adversary would ever take an action because they believed that a significant fraction of US nuclear weapons would not work reliably. Similarly, nuclear weapons have been proven more than effective for excessive destruction. The focus should be on reducing their numbers and explosive power.

A prominent example of an SSP project involving many physicists is the National Ignition Facility, which is funded to advance understanding of nuclear-warhead physics in the absence of nuclear testing. It is also used to understand the behavior of materials in the centers of stars and planets, and it might open a new path to fusion energy. But like other aspects of the SSP, it could have counterproductive results in addition to the beneficial ones.

There are questions physicists must ask themselves. For example, does work on laser implosion encourage the proliferation of thermonuclear-weapons science to other nations? More generally, could improved understanding of weapons physics, in the US or elsewhere, inspire proposals to introduce design changes that would encourage nuclear testing in order to verify that the changed designs still work?

A central justification for nuclear weapons is deterrence. The premise is that wars between major powers are prevented by the danger of catastrophic nuclear use. Nuclear weapons surely deter the expansion of local wars such as the one in Ukraine. But it may be too much to credit nuclear deterrence exclusively, as is often done, for the absence of direct war between the leading military powers since World War II. Many other changes in the international system have also contributed, including the formation of supranational organizations such as the European Union and the United Nations, an increase in the number of democracies, and increased global trade and international scientific collaboration.

In any case, nuclear deterrence risks global catastrophe. Including the Cold War confrontations over West Berlin, the Cuban missile crisis, and multiple close calls from launch-on-warning postures, humanity has escaped a nuclear Armageddon by many strokes of luck. Physicists, familiar with instabilities in physical systems, should be explaining the instabilities of the current nuclear postures and how to reduce them—especially incentives for first use.

In the early years of the nuclear age, eminent physicists working on nuclear weapons in the national labs struggled over the ethics of their work. Hans Bethe once called the hydrogen bomb "the greatest menace to civilization."[5] He later explained his decision to work on it nonetheless: "If I didn't work on the bomb somebody else would—and I had the thought if I were around Los Alamos I might still be a force for disarmament. So I agreed to join in developing the H-bomb. It seemed quite logical. But sometimes I wish I were more consistent an idealist."[6]

Such self-questioning within the physics community, including within nuclear weapons laboratories, currently seems muted. It is time for the renewal of vigorous discussions of how to reduce the dangers from nuclear weapons and of the consequences of research on nuclear weapons. To suggest one route for participation, we invite physical scientists to join the Physicists Coalition for Nuclear Threat Reduction (https://physicistscoalition.org), which the two of us recently cofounded with others. Physicists must act now, for the sake of everyone.

References

1. M. Roser, B. Herre, J. Hassel, *Nuclear Weapons*, Our World in Data (2013).
2. H. M. Kristensen, M. Korda, *Nuclear Notebook*, *Bulletin of the Atomic Scientists* (2023).
3. H. M. Kristensen, *Alert Status of Nuclear Weapons*, Federation of American Scientists (2017).
4. L. Xia et al., *Nat. Food* **3**, 586 (2022).
5. H. A. Bethe, *Sci. Am.* **182**, 18 (1950).
6. S. S. Schweber, *In the Shadow of the Bomb: Oppenheimer, Bethe, and the Moral Responsibility of the Scientist*, Princeton U. Press (2000), p. 166.

Stewart Prager
(sprager@princeton.edu)
Frank N. von Hippel
(fvhippel@princeton.edu)
Princeton University
Princeton, New Jersey

CONTACT PHYSICS TODAY

Letters and commentary are encouraged and should be sent by email to **ptletters@aip.org** (using your surname as the Subject line), or by standard mail to Letters, PHYSICS TODAY, American Center for Physics, One Physics Ellipse, College Park, MD 20740-3842. Please include your name, work affiliation, mailing address, email address, and daytime phone number on your letter and attachments. You can also contact us online at **https://contact.physicstoday.org**. We reserve the right to edit submissions.

Revisiting science and colonialism

Being a history buff, I have read all sorts of "justifications" for colonialism, including Niall Ferguson's book *Empire*, in which he claims British imperialism modernized the world,[1] and Bruce Gilley's controversial article "The case for colonialism," in which he presents a full-throated justification for the practice.[2] But the commentary by Suman Seth (PHYSICS TODAY, December 2022, page 10) is the first piece that I've read that seems to glorify colonialism by linking it to scientific advances.

Seth states, "It is hard to imagine what much of modern science would have looked like without colonialism." Such a statement should be accompanied by a mention of the fact that under colonialism hundreds of millions of people lost their lives, many colonies were looted, and slavery flourished—the consequences of which we still live with today.

Are we supposed to look more fondly on colonialism because some scientific advances may have been delayed a bit in its absence? Before considering where science would be without colonialism, one should consider colonialism's devastating impacts. Colonialism killed more than 50 million native people in the

Americas.[3] King Leopold II's rule over the Congo Free State is associated with the deaths of 5 million to 10 million people, though it was possibly many more.[4] Under British rule, India experienced 165 million excess deaths between 1880 and 1920.[5] An estimated 125 000 to 400 000 civilians died in the First Indochina War, which was fought for liberation from France. Algerian sources say 1.5 million lives were lost in that country's war of liberation against France. And such statistics do not even reflect the cultural genocide that occurred.

It is shameful to imply colonialism was justified for any reason—and that includes its connection to scientific advances.

References

1. N. Ferguson, *Empire: How Britain Made the Modern World*, Allen Lane (2003).
2. B. Gilley, *Acad. Quest.* **31**, 167 (2018); originally published online by *Third World Q.* in September 2017 but withdrawn shortly thereafter following calls for its retraction.
3. A. Koch et al., *Quat. Sci. Rev.* **207**, 13 (2019).
4. J. Gunther, *Inside Africa*, Harper & Brothers (1955); A. Hochschild, *King Leopold's Ghost: A Story of Greed, Terror, and Heroism in Colonial Africa*, Houghton Mifflin (1998).
5. D. Sullivan, J. Hickel, *World Dev.* **161**, 106026 (2023).

Muhammad Sahimi
(moe@usc.edu)
University of Southern California
Los Angeles

~~~

In his December 2022 commentary (page 10), Suman Seth reflects on the historical interconnection between scientific development and colonialism. A fascinating document whose mere existence illuminates that relationship in the 19th-century British empire is *A Manual of Scientific Enquiry; Prepared for the Use of Her Majesty's Navy: and Adapted for Travellers in General* (1849). The book was edited by astronomer John Herschel and can now be found online. It includes sections by such notable scientists as Charles Darwin, who writes on geology, and George Airy, who discusses astronomy. Among the other topics it covers are ethnology, statistics, and magnetism. Armed with its guidance, the officers of the empire could make themselves scientifically useful while ranging the globe.

Ralph Lorenz
(ralph.lorenz@jhuapl.edu)
Johns Hopkins Applied Physics Laboratory
Laurel, Maryland

► **Seth replies:** I am afraid that Muhammad Sahimi has rather misconstrued the point of my commentary. My aim was not to defend, let alone "glorify," the evils of colonialism by linking it with science, but rather to better characterize the nature of modern scientific work. Despite decades of scholarship, many people persist in seeing science—or, at least, "good" science—as a largely apolitical and decontextualized endeavor. Making obvious the fact that modern science could not have existed without connections to multiple devastating colonial projects and that those colonial projects often rested on scientific advancements seemed to be a straightforward way to refute that belief.

Suman Seth
(ss536@cornell.edu)
Cornell University
Ithaca, New York

Malaysian physics in the 1970s

THE AUTHOR (left) and Allen Brailey on vacation in Bali during their stints in the Peace Corps teaching physics in Malaysia. Brailey is holding a copy of *Gravitation* by Charles Misner, Kip Thorne, and John Wheeler.

It was wonderful to see an article about physics in Malaysia in the February issue of Physics Today (page 32). I taught physics at the Universiti Kebangsaan Malaysia (UKM), the National University of Malaysia, from 1975 to 1978 as a Peace Corps volunteer with a master's degree. It was one of the best experiences of my life.

I had many great colleagues in the UKM physics department. Professors Yatim and Lim were particularly memorable. I wish I had a photo of the group like the one of the University of Malaya physics department published in the February feature.

During my time at UKM, I was the first to obtain a grant for a telescope at the school. It was installed on the top of the science building, and many students enjoyed superb views of the Moon, which plays an important role in Islam. I also established an astronomy *istilah* (glossary) by using an algorithm to translate technical terms from English into Malay, which I learned during Peace Corps training in Kuantan, on the South China Sea. I also taught beginning-level astronomy classes in the language. It was a great experience.

For the more advanced courses—nuclear physics and graduate-level electricity and magnetism—my students knew English much better than I knew Malay, which was helpful. I must say that the students in those courses were fantastic: Each of them always turned the homework in on time and had excellent handwriting. They spoiled me into imagining that being a professor in the US would be similarly easy!

In one fun anecdote from my time teaching at Reed College, a colleague of mine, David Griffiths, could hardly believe that I taught electricity and magnetism—out of John David Jackson's *Classical Electrodynamics*, no less—in Malaysia and in Malay. Who could blame him? But David, himself the author of a popular undergraduate textbook on the subject, was convinced when I showed him my lecture notes.

I am eternally thankful to the US Peace Corps and UKM for three spectacular years in a great part of the world.

Johnny Powell
(dna@reed.edu)
Reed College
Portland, Oregon

Correction

April 2023, page 49—The review incorrectly characterized Arnold Sommerfeld as an experimental physicist. He was primarily a theorist. PT
~~~

Uranus's hidden polar cyclone, revealed

Microwave observations peer into the atmospheric dynamics of the oddball seventh planet.

Jupiter has *Juno*. Saturn had *Cassini*. Even tiny Pluto had *New Horizons*. But when it comes to exploration, Uranus and Neptune, the solar system's ice giants, have been left out in the cold. Neither has been visited by any spacecraft since (or before) *Voyager 2*, which flew by the former in 1986 and the latter in 1989.

Voyager 2 captured stunning photographs of the outer planets: Jupiter's mottled stormy surface, Saturn's elegant stripes and groovy rings, and Neptune's mesmerizing swirls of blue. Its images of Uranus, in contrast, weren't much to look at. As shown in figure 1 at left, the seventh planet appeared as a dull, pale sphere, as plain and featureless as a Ping-Pong ball.

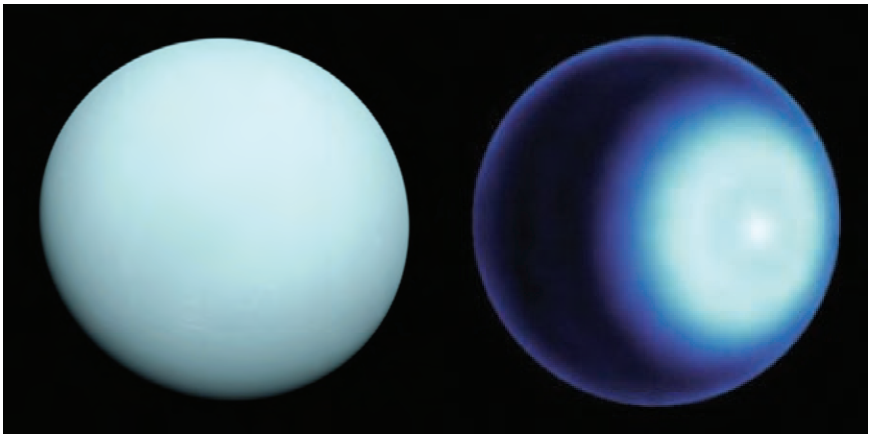


FIGURE 1. PALE BLUE DOT. When *Voyager 2* flew by Uranus in 1986, it sent back the visible-light photograph, shown here at left, of the planet's outer layer of methane clouds. In the decades since then, Earth-based observations at longer wavelengths have probed deeper and uncovered the planet's layers of hidden complexity. Shown at right is a new microwave image with signs of a cyclone at the Uranian north pole. (Left panel courtesy of NASA/JPL-Caltech; right panel courtesy of NASA/JPL-Caltech/VLA.)

But "boring," it turns out, is only skin deep. *Voyager 2*'s visible-light images show Uranus's surface layer of methane clouds. In the decades since the flyby, researchers have looked deeper into the planet by imaging it at longer wavelengths, and they've discovered that beneath the plain exterior lurks a complex and dynamic atmosphere. Now Alex Akins, of NASA's Jet Propulsion Laboratory, and colleagues have used the Very Large Array—a squadron of 28 radio dishes (one is a spare) spread over tens of kilometers on the plains of western New Mexico—to image Uranus at wavelengths from 0.7 cm to 5 cm. And they've uncovered a clear sign of a cyclone circling the planet's north pole.[1]

Polar cyclones exist on Earth—and on every other planet with an atmosphere in the solar system. But Uranus is unusual enough that it wasn't a given that one would be seen there. "In comparative planetology, we're interested in what changes in planetary systems under different conditions and what stays the same," says Akins. "So the observation of polar cyclones on every planet with a substantial atmosphere tells us that they're likely present on much of the exoplanet population as well."

Sideways seasons

Akins and colleagues' microwave images, including the one in figure 1 at right, show the Uranian north pole close to head-on. That's possible with an Earth-based telescope only because Uranus has an exceptional orbital geometry. Unlike most other planets, whose rotational axes are at least roughly parallel to their orbital axes, Uranus looks like it's been knocked on its side. During its 84-Earth-year orbit, its poles take turns pointing straight toward the Sun—and toward Earth.

The orbital configuration also potentially affects the Uranian seasons and atmospheric dynamics. On Earth and similar planets, the poles are colder than the equator because they get less direct sunlight. In response to the thermal imbalance, hot air rises at the equator, drifts to higher latitudes, and sinks. When that convection cycle interacts with Earth's rotation, it gives rise to the pattern of pressures and prevailing winds, including the Arctic and Antarctic polar vortices, that dominates our local weather and climate. (See the article by Thomas Birner, Sean Davis, and Dian Seidel, Physics Today, December 2014, page 38.)

On Uranus, however, everything is different. Averaged over a year, the poles get more sunlight than the equator, not less. On the other hand, Uranus is 20 times as distant from the Sun as Earth is, so it gets barely 1/400 the sunlight. That may not even be enough to drive significant seasonal or regional variation.

The most recent Uranian equinox was in 2007. The current northern-hemisphere spring is the first time the north pole has been visible from Earth since 1965—and it's the first chance astronomers have had to get a good look at either pole since the south pole faded from view in the early 2000s.

Microwave-astronomy capabilities have improved a lot during those decades, including at the Very Large Array (shown in figure 2), which underwent a significant upgrade[2] that was completed in 2012. Akins and colleagues' new images show the north polar region in more detail than was possible before, and they're newly able to resolve the polar cyclone: the small bright spot right

FIGURE 2. RADIO FORMATION. The Very Large Array, a roughly two-hour drive outside of Albuquerque, New Mexico, is one of the world's foremost observatories for radio- and microwave-frequency astronomy. With all of its radio dishes aimed at the same spot on the sky, the array can capture images in exquisite detail, even resolving features on our solar system's outer planets. (Courtesy of T. Burchell, NRAO/AUI/NSF.)

at the pole encircled by a faintly darker collar at about 80° N. Those features, the researchers conclude, stem from the temperature–pressure pattern at the center of a cyclone, similar to the eye of a terrestrial hurricane.

Around again

Not everything in the new images is new. The Very Large Array has been fully operational since 1980, albeit with lower bandwidth and resolution than in recent years, and microwave observations of Uranus date back nearly that far—close to half a Uranian year. And they've consistently shown the same feature that dominates Akins and colleagues' images and constitutes one of the many mysteries of the Uranian atmosphere's circulation: The broad polar regions, with latitudes higher than about 45° north or south, glow brighter than the equatorial zone.

Microwave imaging records thermal radiation, not reflected light. So one reason the Uranian poles might be brighter than the equator is that they're hotter. They do get more sunlight, after all. But that explanation is less than satisfying because the brightness contrast persists regardless of the Uranian season, even for a pole just emerging from the dead of winter. And the temperature difference required to produce the contrast—several tens of kelvin—would be hard to explain in any circumstance.

The polar brightness could also be the result of a difference in chemical composition. Uranus is known as an ice giant because of its richness in ice-forming materials—such as water, ammonia, and methane, as opposed to the hydrogen and helium that dominate the gas giants Jupiter and Saturn—not because those substances are necessarily present in their solid form. The chemical diversity gives the atmosphere a complex layered structure, with water clouds at the bottom and methane clouds at the top. But the layers might not be the same everywhere.

Methane doesn't strongly absorb microwaves, but gases deeper in the atmosphere, such as ammonia and hydrogen sulfide, do. If those low-lying absorbers somehow got churned up to higher altitudes near the equator, they could block the microwave emissions and make that region appear darker. "The observations are consistent with a simple model of atmospheric circulation where air rises at lower latitudes and descends at higher latitudes," says Akins. Such a circulation pattern is familiar on Earth, with its warm equator and cold poles, but how the same phenomenon could arise on Uranus is less clear. "We're not exactly sure what causes it, especially since insolation is so weak," he adds. "I hope more folks in the community get excited about Uranus and can help provide some answers."

Another voyage

As the northern Uranian spring progresses into summer, keeping an eye on the polar cyclone could offer valuable new insights into how a polar atmosphere behaves after 42 continuous years in the dark—and what seasonal effects, if any, are present in Uranus's atmosphere. Although it's too soon to tell for sure, Akins and colleagues have already seen hints that the cyclone may have strengthened a bit during its short time in the Sun.

Ultimately, though, Earth-based observations can only do so much. They're limited in their resolution and spectral bandwidth, and they can't reach the meter-long wavelengths needed to image Uranus's deep water clouds. To really unravel the mysteries of the Uranian atmosphere, it will probably be necessary to get a close-up view once again.

Happily for seventh-planet aficionados, a Uranus orbiter and probe was deemed the top-priority flagship mission by the latest planetary science decadal survey, released in April 2022 by the National Academies of Sciences, Engineering, and Medicine. The recommendation alone is no guarantee that the project will come to fruition, but it bodes well. The top two priorities of the previous survey—a Mars sample-return mission and the *Europa Clipper*—were both funded, with launches planned in the next few years.

For Uranian science, time is of the essence. If launched in 2031 or 2032, a Uranus-bound spacecraft could capitalize on a gravity assist from Jupiter and reach its destination in a mere 13 years. If launched later, its journey will take much longer.

Johanna Miller

References

1. A. Akins et al., *Geophys. Res. Lett.* **50**, e2023GL102872 (2023).
2. R. A. Perley et al., *Astrophys. J. Lett.* **739**, L1 (2011).

Understanding how metal-nitride ferroelectrics switch their polarization

Transmission electron microscopy images and first-principles calculations suggest that the atoms adopt a disordered but low-energy configuration that facilitates the switching.

Studied for more than a century, ferroelectric materials exhibit a spontaneous polarization in one or more directions along a crystal axis. Thermodynamically stable, the polarized states can be switched from one to the other by applying an electric field that exceeds what's known as the coercive field E_c. That switchability provides the basis for the nonvolatile RAMs in computing. (See the article by Orlando Auciello, James F. Scott, and Ramamoorthy Ramesh, Physics Today, July 1998, page 22.)

Since the 1960s, electrical engineers have been designing memory elements based on conventional ferroelectrics such as the perovskite barium titanate. But manufacture is complicated by the challenge of integrating those materials with silicon-based semiconductors. What's more, the memory elements are difficult to scale down to atomic dimensions for energy efficiency. Between 2019 and 2021, researchers discovered that crystalline films of alloyed aluminum nitride are ferroelectrics that could solve both problems. Boron-doped AlN, in particular, is easy to integrate, as it consists exclusively of elements common in silicon electronics.

The discovery was a surprise to most scientists. The films were well-known pyroelectric and piezoelectric crystals, but few believed they could be ferroelectric, because their coercive fields are just too perilously close to the field at which the materials experience dielectric breakdown. Apply a high enough electric field to switch the polarization and you risk destroying the material.

Pennsylvania State University materials scientists Jon-Paul Maria, Susan Trolier-McKinstry, and Ismaila Dabo, who had demonstrated ferroelectricity in B-doped AlN two years ago,[1] have now teamed up with Carnegie Mellon University materials scientists Sebastian Calderon and Elizabeth Dickey to address the dielectric-breakdown problem.[2] The collaborators realized that if they understood the mechanism by which the polarization switches at the atomic scale, they could manipulate it—for example, by straining the film, growing it thinner, or altering its dopant concentration.

Such tricks may dramatically lower the coercive field from roughly 4–6 MV/cm in metal nitrides down to levels approaching 1 MV/cm. That transformation would render the new ferroelectrics more practical for applications in memory, energy-harvesting, and high-speed and high-power circuits.

Double duty

Fortunately, there are a few ways to ensure that the coercive field is lower than the dielectric breakdown field. At elevated temperatures, the margin separating E_c and the breakdown field increases, for instance.

As deposited on a tungsten electrode, $Al_{15/16}B_{1/16}N$ films grow with a polarization that points downward (into the substrate)—a configuration shown in figure 1a and referred to as N-polar growth. The image depicts the thickness of a film, grown upward from the bottom of the frame. The tetrahedron diagrams show the positions of the metal atoms Al or B (gray) relative to the N atoms (blue). The metal atom in each tetrahedron is bonded to four N atoms.

To experimentally study the films' local structure and polarization, the researchers used scanning transmission electron microscopy (TEM). And in a standard imaging mode known as differ-

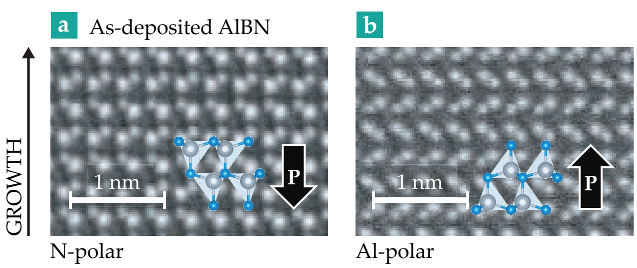


FIGURE 1. BEFORE AND AFTER polarization reversal. These transmission electron microscopy images are on-edge views of a boron-doped aluminum nitride film—a mere 6 nm thick—that show the projections of N, Al, and B atoms through the lattice. **(a)** In AlBN as it's deposited on tungsten, the polarization orientation **P** is down, or N-polar. **(b)** After the surface expels enough charge to exceed the coercive field, the polarization flips upward (Al-polar) and the angles between the N atoms and the Al or B atoms change their relative orientation. (Adapted from ref. 2.)

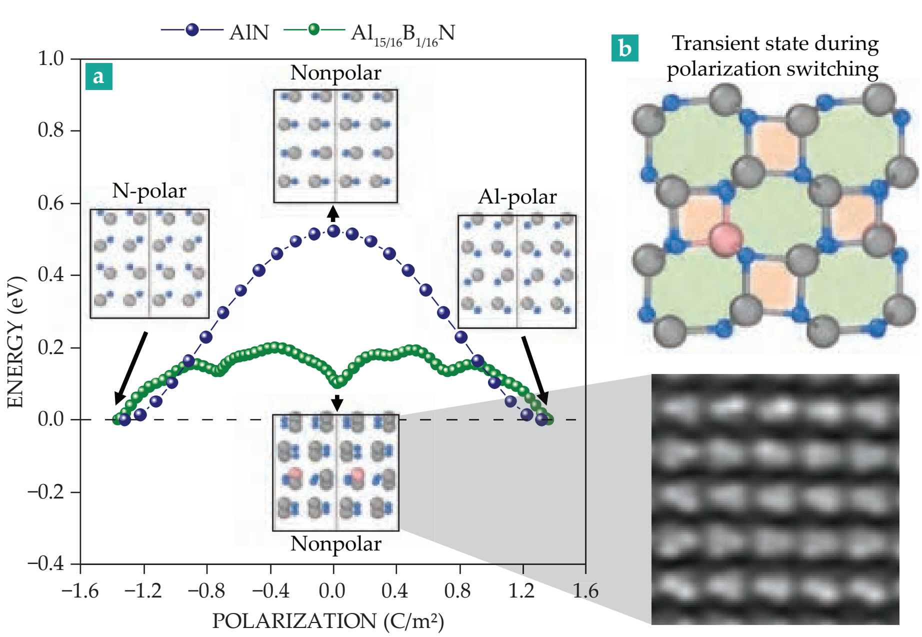


FIGURE 2. A FIRST-PRINCIPLES CALCULATION. Researchers move aluminum and nitrogen atoms in a film of aluminum nitride to determine the minimum energy pathway required to switch its polarization from N-polar to Al-polar. **(a)** Along one proposed path (blue), Al (gray) and N (light blue) atoms in each tetrahedron of a pure AlN film adopt a nonpolar, linear configuration midway. But that configuration, which has a 0.5 eV energy barrier, is never observed by transmission electron microscopy. Along an alternative path (green), the atoms switch polarization in multiple stages. **(b)** The green path's intermediate structure was confirmed (bottom) in an experiment and predicted (top) in a nearly 30-year-old simulation to include area defects in the form of four- and eight-membered rings.[3] (Adapted from ref. 2.)

ential phase contrast, the microscope images the sublattices of the heavier (Al) and lighter (N) atoms at the same time. The atomic columns appear as distinct projections through the lattice.

The TEM beam also can trigger a polarization switch: As electrons scatter in the film, they ionize local Al, B, and N atoms. Electrons are ejected, and the sample becomes positively charged. When a local electric field exceeds E_c, the film's polarization switches upward (figure 1b).

Imaging the configuration of atoms during the switching process takes about seven minutes. But that time is largely spent charging the sample to the coercive field. And the TEM affords researchers the luxury of imaging continually while the charging takes place. In practice, actual ferroelectric devices flip their polarization far faster—on time scales of about a nanosecond.

Inversion through a flattened plane

When they joined the project, Maria and Dickey initially imagined that each metal atom in a tetrahedron would break one of its N bonds. The Al or B atom would then move into the plane of the three remaining N atoms before protruding to the other side, where it would then bond with another N atom. That would be the simplest, most intuitive way to switch the polarization.

But for it to work, they realized, B atoms would be needed to flatten the tetrahedra. Like graphite, BN crystallizes in a planar hexagonal lattice, and B is energetically more likely to adopt a threefold coordination in covalently bonded sheets than to adopt a four-fold one.

By passing through an intermediate stage in which the AlN tetrahedra flatten, the Al–N dipoles were expected to cancel and the polarization to drop to zero. In the collaboration's new work, first-principles calculations by graduate student Steven Baksa from the Dabo group bore out that imagined scenario (the blue curve in figure 2a), but only in the case of an undoped film of pure AlN. In the calculation, the 32 Al and 32 N atoms are incrementally nudged along a proposed polarization path, during which the crystal structure diverges from a thermodynamic minimum energy state into a higher-energy, zero-polarization state. But the 0.5 eV energy barrier found by the calculation prevents that nonpolar configuration from forming.

When the researchers introduced a single B atom into the calculation, a more complicated atomic structure emerged instead—one that includes substantial local bonding and structural distortions. Both N-polar and Al-polar bonding configurations appear in the transient atomic structure, combining both upward and downward polarizations in a doubled unit cell. The B-doped AlN calculation revealed that the lowest energy path to switch the polarization is not the smooth parabolic one, but rather one with local minima (green curve in figure 2a).

Reassuringly, the collaboration's TEM image of the B-doped film (figure 2b, bottom) experimentally confirms that more complicated structure. Still more reassurance came when the researchers recognized the structure as a form of gallium nitride (figure 2b, top). It contains four- and eight-membered rings—crystallographically distinct from the parent lattice and predicted in the literature nearly 30 years ago[3]—in place of some of the hexagonal rings that otherwise populate the film.

The formation energy of those intermediate states—just 0.2 eV—is small enough that making them is achievable with an electric field that doesn't destroy the material. Is there an even better switching pathway that would lower the coercive field while retaining the temperature-stable polarization? The researchers don't yet know, but Trolier-McKinstry expresses hopeful enthusiasm: "We're really looking forward to finding out. For starters, we will be exploring compositional modifications that could result in lower local energies for the reversal process."

Mark Wilson

References

1. J. Hayden et al., *Phys. Rev. Mater.* **5**, 04412 (2021).
2. S. Calderon V et al., *Science* **380**, 1034 (2023).
3. J. E. Northrup, J. Neugebauer, L. T. Romano, *Phys. Rev. Lett.* **77**, 103 (1996). PT

ITER appears unstoppable despite recent setbacks

Repairs could take up to two years, but project officials believe they can perform them in parallel with the machine's assembly. Regulatory concerns are unresolved.

ITER, the international project to build a giant tokamak to achieve a burning fusion plasma, already faced years of delays and cost increases even before defective components were discovered last year. But project officials say they likely can't provide an estimate of the length of the delays or how much more ITER will cost until the end of 2024.

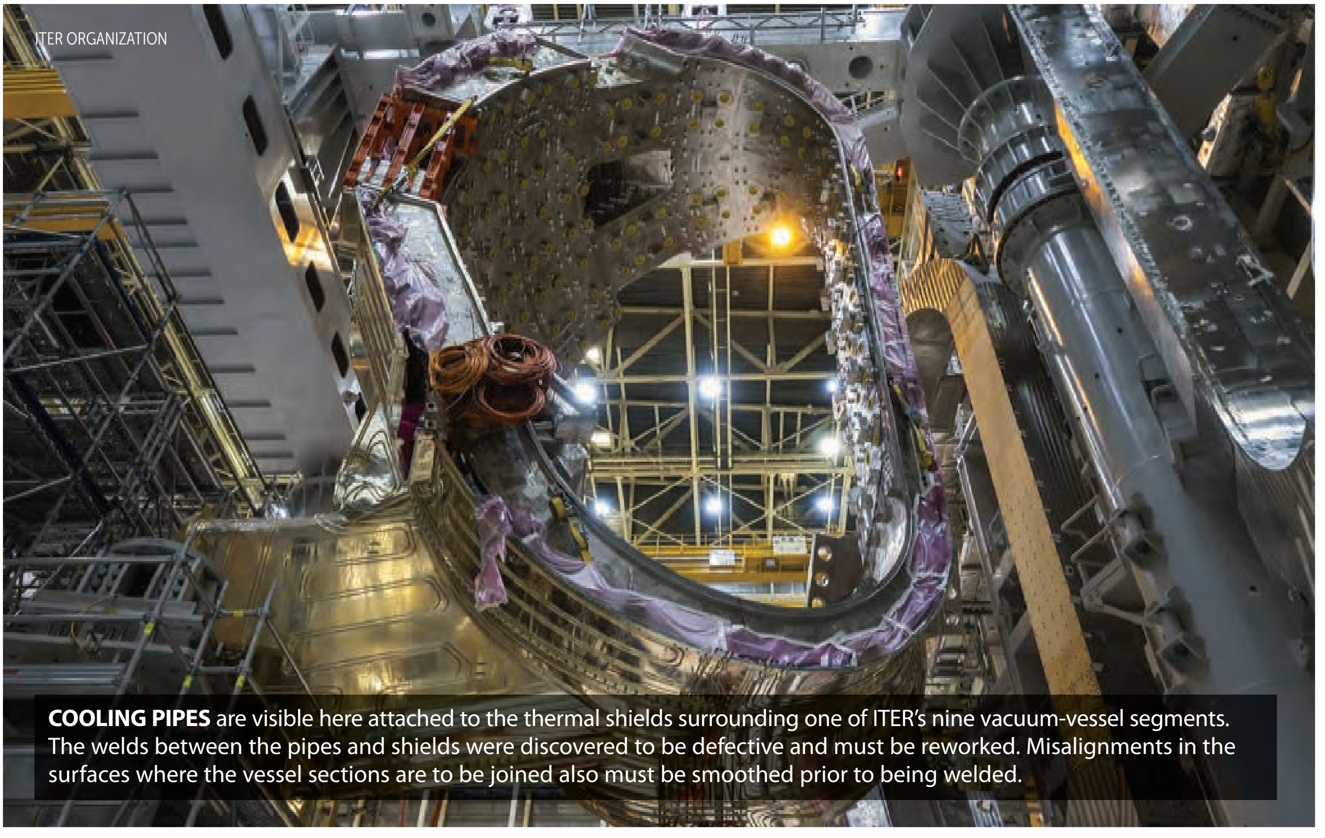
Some internal estimates have indicated ITER's completion could be delayed by as much as 35 months from 2025, the date for the machine's commissioning envisioned in the project's 2016 baseline, says ITER spokesperson Laban Coblentz. Shorter delay estimates also have been discussed internally, he says, but "none of the numbers are official or reliable." Project leaders first warned in 2020 that the 2025 start date is not achievable.

Coblentz insists that no official estimate can be provided until ITER management produces a revised baseline—and the ITER Council, ITER's governing body, approves it. "Every journalist—and every stakeholder—would prefer to have a precise answer about the new expected schedule," says Coblentz. "But the evaluation is complex, and the easy choice is to cite selective elements and extrapolate conclusions."

In April 2022, cracks were found in some of the 23 km of pipes that will conduct cooling water through the thermal shields sandwiched between ITER's vacuum chamber and the surrounding superconducting toroidal-field magnet coils. The defects were traced to inadequate surface preparation prior to welding the pipes to the shields. Most of the shields have been delivered to the ITER site at Cadarache, in southern France. The one vacuum chamber section that's been installed in the reactor so far was removed in July and is being taken apart to allow rework of the faulty components.

A second major defect, misalignments in the welding surfaces of the four vacuum-chamber segments manufactured in South Korea, was discovered in 2020. (See PHYSICS TODAY, May 2022, page 20.) The surfaces must be smoothed out—voids filled in and high points ground off—before they can be welded together. The remaining five segments of the doughnut-shaped vacuum chamber are still being manufactured in the European Union (EU).

Repair work on the two defects began



ITER ORGANIZATION

COOLING PIPES are visible here attached to the thermal shields surrounding one of ITER's nine vacuum-vessel segments. The welds between the pipes and shields were discovered to be defective and must be reworked. Misalignments in the surfaces where the vessel sections are to be joined also must be smoothed prior to being welded.

in July and is expected to take two years to complete, says Coblentz. But he and Tim Luce, ITER's head of science and operation, say the work can be done in parallel with the machine's assembly and won't necessarily further delay the project schedule. "We don't need all nine sectors to begin to assemble the vacuum vessel. We need the first three," says Luce. The assembly sequence can proceed as the others continue to be repaired.

Many causes for delays

Much of the yet-to-be-quantified delay is attributable to the COVID-19 pandemic and related supply-chain issues. Technical challenges tied to first-of-a-kind components with multiyear fabrication timelines, such as the magnets and vacuum-vessel sectors, also played a role, says Coblentz.

Another delay will come from testing components to offset future risks. For example, revised plans call for testing the toroidal field coils at 4 K in the completed cryogenics plant prior to their installation. Such testing wasn't specified in the original baseline.

A proposed schedule that was reviewed and rejected last year by the ITER Council would have been immediately outdated with the discovery of the manufacturing defects, says Luce.

ITER director general Pietro Barabaschi, who declined an interview request, acknowledged in news releases that the cost of the repairs "will not be insubstantial." Barabaschi took over following the death of Bernard Bigot last year.

ITER's schedule will also be affected by the French Nuclear Safety Authority (ASN), whose February 2022 order to halt assembly remains in effect. The agency has questioned the adequacy of ITER's radiological shielding, and it worries that adding on to the 3-m-thick concrete shielding that already surrounds the reactor pit would raise the mass of the reactor beyond the capacity of its support system. The ASN also expressed concern over the vacuum-vessel welds. ITER officials had hoped the ASN would lift its hold last fall, but an ASN spokesperson said in late June that ITER had yet to satisfactorily respond to the regulatory issues.

As part of the re-baselining exercise, ITER management is planning to compress the previous timetable for the onset of deuterium-only experiments. The goal would be to adhere as closely as possible to the previous 2035 target date for the onset of tritium experiments. The current baseline, says Luce, calls for a "first plasma" upon completion of construction, to ensure that the vacuum vessel, magnets, and other physical plant components function

ITER'S LIFE began during a 1985 summit of US president Ronald Reagan and Soviet leader Mikhail Gorbachev. Following years of design efforts and negotiations over the location, a site in France was selected and construction began in 2010. Completion, originally planned for 2025, will be delayed, officials acknowledge, but a new timetable won't be available until next year. (All images courtesy of ITER Organization.)

What's old is new in DOE's choice of fusion hopefuls

Among the eight fusion startup companies that will share $46 million in grants to help build commercial power plants are four that are pursuing approaches—stellarator, magnetic mirror, and Z pinch—that were once major programs in Department of Energy labs before being mostly abandoned in favor of tokamaks.

Awardees also include two companies that are pursuing differing approaches to inertial fusion. Two others are developing tokamaks.

DOE says the grants will assist the companies to develop their respective designs for commercial fusion pilot plants in 5–10 years. At least one recipient, the MIT spin-off Commonwealth Fusion Systems, has said it will begin constructing a pilot power plant within five years. It has raised more than $2 billion. The well-established Cambridge University spin-off Tokamak Energy has been working on its spherical tokamak design for more than a decade in the UK. Officially, its award went to a US subsidiary in West Virginia.

The 18-month grants are the first tranche of an anticipated $415 million in DOE support over five years. But awardees won't receive any of the money unless they produce acceptable "pre-conceptual designs and technology roadmaps" for their power plants within the next year and a half. According to DOE a preconceptual design is similar to a conceptual design but at lower levels of fidelity and with greater uncertainties. A technology road map details the required critical-path R&D, including any intermediate test facilities, required for a particular pilot plant conceptual design.

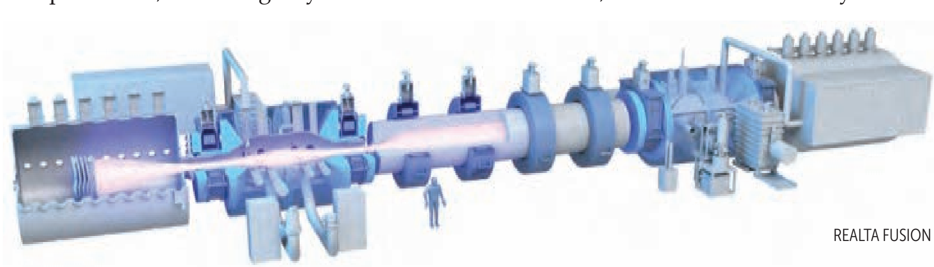
Thea Energy of New Jersey, and Type One Energy Group, based in Madison, Wisconsin, each are betting on stellarators, a magnetic-confinement concept that Lyman Spitzer pursued when he founded in the 1950s what would become the Princeton Plasma Physics Laboratory. Like the tokamak, a stellarator creates a doughnut-shaped plasma, but it does so with a twist that requires a complex configuration of magnets to maintain.

Another Madison-based firm, Realta Fusion, is counting on magnetic mirrors, which was a major DOE program in the 1970s and early 1980s. In those cylindrical-shaped devices, charged plasma particles trapped in a magnetic field hit a point along the field line where they reverse direction. Some will collide with others and fuse as they bounce back and forth.

Zap Energy, in Everett, Washington, hopes to create commercial fusion using Z pinch, a technology that Los Alamos National Laboratory explored with its Perhapsatron machine in the 1950s. In a Z pinch, a thin line of plasma is magnetically confined and compressed by an electrical current running through the plasma. One version of Z pinch is the massive Z machine at Sandia National Laboratories, which is primarily used in support of nuclear weapons. By comparison, Zap's reactor would fit on a tabletop.

In recent decades, DOE's Fusion Energy Sciences program has focused nearly exclusively on tokamaks. But it has continued to support small-scale nonmainstream concepts. The Advanced Research Projects Agency–Energy in particular has backed academic and privately funded research on non-tokamak approaches in recent years. In a news release, the agency boasted that the milestone awards validated its own technology choices.

Focused Energy, located in Austin, Texas, and Darmstadt,



REALTA FUSION

A RENDERING of Realta Fusion's conceptual tandem-mirror reactor, consisting of two end cells on either side of a longer central cell in which most of the fusion will occur.

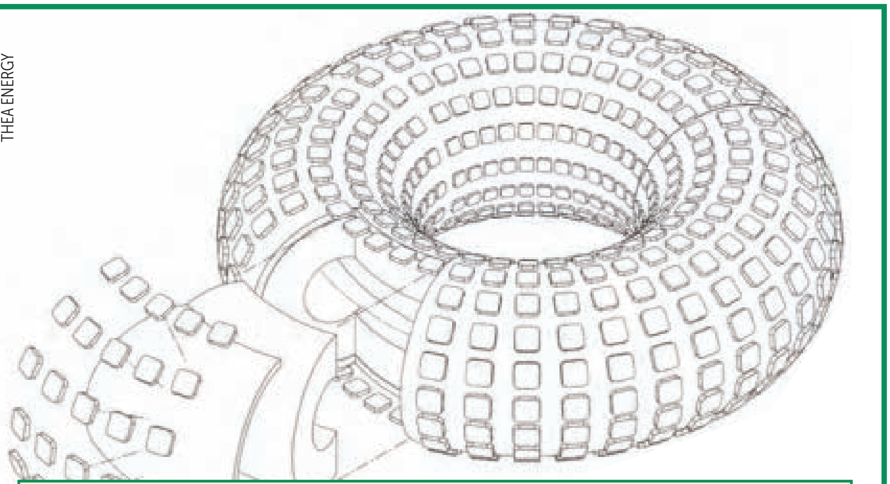
properly. A two-year pause was then planned to permit the installation of remaining vacuum-vessel components and additional heating systems. Only then would experiments with deuterium begin.

Project officials have decided to jettison the plan to coat the vacuum chamber's walls with beryllium. While workforce exposure to the toxic metal contributed to the decision to replace it with tungsten, Luce says the major reason for the design change is to increase ITER's relevance to future commercial fusion power plants. Tungsten is expected to better withstand the constant bombardment by high-energy fusion neutrons.

Costs are a mystery

As the host ITER partner, the EU contributes 45% of ITER's cost. The other ITER partners—China, India, Japan, Russia, South Korea, and the US—each contribute 9%. Because the seven partners each have different labor, materials, and other expenses associated with their contributions, which are mainly in the form of fabricated reactor components, the project's exact cost in dollars or euros may never be known. Indeed, ITER has its own currency, called ITER Units of Account.

At the request of the House Science, Space, and Technology Committee, however, Bigot in 2017 estimated ITER's cost would total $25 billion through 2035, when tritium experiments were then supposed to begin. But the EU that same year estimated its share of the project alone would total €18.1 billion ($19.6 billion) through 2035. By extrapolation, the total ITER cost during that period would be €41 billion if the entire project were to be undertaken in the EU.



THEA ENERGY

THEA ENERGY'S stellarator design consists of arrays of small magnet coils, seen here as the outer layer of squares. Other stellarators have used thicker and more complex three-dimensional coils. The coils will be held in place by a support structure, which surrounds a thicker layer, called the blanket, encompassing the plasma.

Germany, was another award recipient, as was Xcimer Energy, in Redwood City, California. They are pursuing different laser pathways to inertial fusion (see Physics Today, March 2023, page 25).

The awards provide an imprimatur of sorts from DOE. "The amount is less important than being part of the program," says Benj Conway, CEO of Zap Energy, which stands to get $5 million. "Being selected as part of the program reflects our progress, the credibility of our approach, what we've managed to achieve, and the plan going forward." Zap, which has raised over $200 million, has commissioned an experimental device that Conway says he expects to achieve breakeven—when fusion energy produced equals the input energy—within a year.

Assuming they meet their milestones, awardees will be eligible for follow-on grants for up to five years. The $415 million Congress has authorized for the program is subject to annual appropriations. The program's success will depend on receiving the full out-year funding, says Conway. "That's where it will make the real difference to us."

Realta chief technology officer Cary Forest and Thea Energy CEO Brian Berzin say the advent of high-temperature, high-field superconducting magnets has been an important element in their respective development paths. Commonwealth Fusion Systems is supplying the high-temperature superconductor magnets for Realta's next-generation mirror device.

David Kramer

The US Department of Energy in 2018 estimated ITER's cost would be $65 billion if all the work were to be done in the US. That didn't include operating expenses during the 2025–35 period of commissioning and initial experiments. DOE's estimates routinely include a large contingency; ITER's do not. (See "ITER disputes DOE's cost estimate of fusion project," Physics Today online, 16 April 2018.)

The latest ITER setbacks were shrugged off during a 13 June hearing by the House science committee devoted to fusion. Kathryn McCarthy, director of the US ITER Project office at Oak Ridge National Laboratory, mentioned the need for repairs in her testimony, but none of the lawmakers who were present pursued the topic. McCarthy noted that the problematic components were not made in the US.

Former representative Jerry McNerney (D-CA) caused a stir last year when, after a visit to ITER, he told his fellow science committee members he was informed that the defective components could be "project-ending." Luce, who met with McNerney, says that was a misunderstanding: a project official told

McNerney that it would have been extremely difficult to fix the shields had the defects not been uncovered until the reactor had been assembled.

No one has called, at least publicly, for ITER to be abandoned. Luce says he's seen no signs of any of the partners defecting. "In fact there have been positive signs," he says, one of which is that the US and India have recently paid their contribution arrears.

At the June semiannual meeting of the ITER Council in Cadarache, member nations "reaffirmed their strong belief in the value of the ITER mission, and resolved to work together to find timely solutions to facilitate ITER's success," according to a communiqué.

Even Robert Hirsch, the former head of DOE's fusion program and a longtime critic of tokamaks, says ITER should continue. "There's no question that there will be some benefit," he says. But because of their complexity, Hirsch predicts that tokamaks will never become a commercially viable energy source. "ITER never should have happened. Having said that, it is happening, and it seems to me that, practically speaking, people can't walk away from it."

Coblentz says that ITER has helped inspire the emergence of private-sector fusion companies. "We are demonstrating that these massive, precise components needed for fusion energy can be built at industrial scale, and we are developing the required new technologies as we go."

David Kramer

Tidal turbine development ebbs and flows

The renewable energy technology can benefit remote coastal communities that want to reduce carbon emissions.

In Canada's Haida Gwaii archipelago, roughly 120 km off the northern coast of British Columbia, the north electrical grid uses more than 7 million liters of diesel to provide power to about 2500 people each year (the south grid uses roughly 3 million liters for roughly 2000 people). For long-term resident Laird Bateham, the predictability of the tides pointed to an obvious alternative. He founded Yourbrook Energy Systems in 2010 to develop a turbine for harvesting power from the local tidal currents. He and his team developed a prototype and are beginning front-end engineering of a project intended to deliver 500 kW of clean and reliable power to the isolated coastal community. That would provide for 20% of the population's average annual use. His colleague Clyde Greenough says, "We want to leave the world a better place by doing our part to slow climate change."

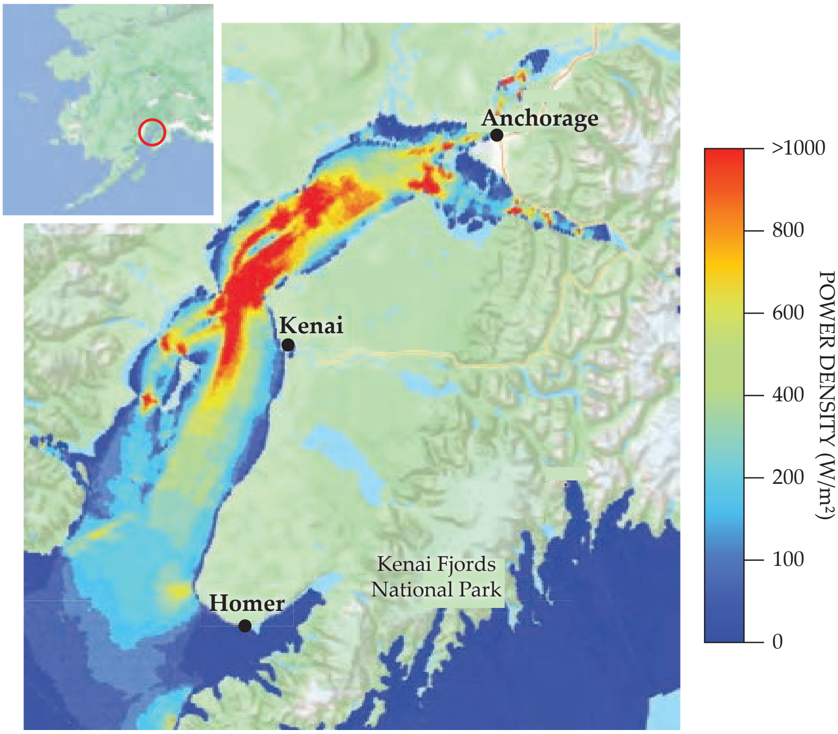
ROWS OF TURBINES in the currents generated by the rise and fall of the tides offer a predictable way of producing electricity. Here, an array of 100 kW turbines sits on the Bluemull Sound seafloor off Scotland's Shetland Islands.

The twice-daily rise and fall of the tides drives powerful, predictable currents when seawater flows toward and away from Earth's coastlines. A turbine placed in the current's path can harness that power. The moving water pushes the turbines' blades, causing them to spin and drive a rotor that powers an electrical generator.

Tidal power has been harvested since the Middle Ages, when people retained incoming tidewater in storage ponds and used the outgoing flow to turn waterwheels for milling grain. In modern times, the world's largest tidal power plants are located in France and South Korea, with 240 MW and 254 MW electricity generation capacity, respectively. There, dam-like structures called barrages span ocean inlets or bays to capture incoming tidal water and generate electricity as the basin fills and empties. The installations can control flows using sluice gates, but because barrages are large scale (making them expensive to construct), they can disrupt the local ecosystem by altering lagoon salinity and animal movements.

In the past decade, motivated by advances in turbine technology and an increased urgency to find energy alternatives to fossil fuels, researchers have worked to generate electrical power by placing small numbers of turbines in strong tidal currents.

Projects like Yourbrook Energy are on



NREL MARINE ENERGY ATLAS

MAPS OF TIDAL RANGES, buoy measurements, and seafloor topographies have led to hydrodynamic models of depth-averaged tidal currents where turbines may be valuable for generating energy. Shown here is the distribution of tidal power density, developed from model simulations, in Alaska's Cook Inlet.

the rise. Tidal turbines could be valuable in helping small coastal communities meet their energy needs. Months-long turbine demonstrations are showing where technical difficulties remain. Scaling up brings challenges related to cost and corrosion and to maintenance of components that are hard to access underwater. The turbines' rotating blades can threaten marine life, and the seafloor platforms that support the turbines can alter flow and induce erosion. In addition, suitable sites are often privately owned, making access difficult. But researchers are coming up with creative technical designs, and turbine developers are finding appropriate sites for testing and installing their technologies.

Predictable, but how practical?

In May 2023, the US Department of Energy announced a $45 million funding opportunity for the development of tidal energy projects. An assessment conducted in 2021 by DOE's National Renewable Energy Laboratory (NREL) determined that the total energy available in the US from tidal currents in the country was 220 TWh/yr. That amount represents roughly 5% of the electricity generated in the US in 2019. The assessment, based on coastal ocean dynamics modeling, estimated that the resource could power 21 million homes. Energy available from all marine resources—including tidal currents, river currents, ocean waves, and ocean thermal gradients—amounts to 2300 TWh/yr, according to the assessment.

The National Research Council in 2013 looked at an earlier set of assessments produced by DOE and raised concerns that are still relevant. For example, the committee that conducted the review suggested that the theoretically available power could be a significant overestimation of the practically available resource. "There's a tendency for proponents to exaggerate the potential and not look at the practical side," says Chris Garrett, an oceanographer emeritus at the University of Victoria and a member of the National Academies review committee. Many evaluations that make a site look appealing for a tidal array based on available resources ignore the impact on marine life and channel erosion. "Things that sound green in principle are green in a small-scale situation," he says. "But even if large scale is feasible, it may not be green."

Garrett's calculations also indicate that proponents overestimate the power potential of tidal turbines. Introducing turbines to a channel slows the water flow. "At some point, adding one more turbine produces less additional power than is lost by the reduced output of the other turbines," Garrett explains. According to a simple study that he and Patrick Cummins published in 2008, the maximum attainable power from an array of turbines in a channel is associated with a 42% reduction in current speed.

Similar studies by researchers across the globe have mapped tidal flow patterns and speeds in specific channels and calculated the power that could be captured by one or more turbines. The sites that can deliver the most power have large tidal ranges, have strong currents, and are located along narrow tidal channels. High-resolution simulations of the sites with the highest tidal power densities can evaluate how energy extraction might affect flow behavior far from a turbine array. Such studies could help to guide site selection and turbine placement.

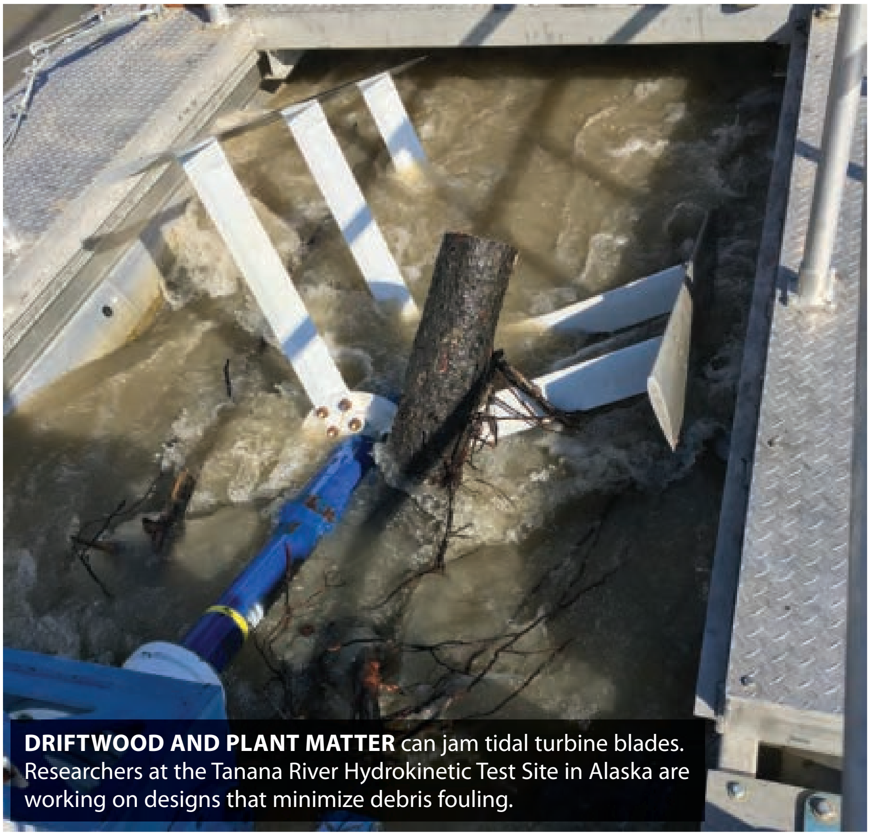
A 2019 DOE overview of marine renewable energy estimated that tidal energy costs \$130–\$280/MWh. An NREL technology baseline report published in 2023 said that in 2021, onshore wind energy cost as little as \$30/MWh, and utility-scale solar \$39/MWh. The cost of any renewable energy technology, however, tends to decrease as the technology improves, mass production increases, and the supporting industry matures.

Where tidal energy excels is its predictability. Unlike wind and solar, tidal currents are "incredibly reliable," says Levi Kilcher, a marine energy researcher at the NREL and coauthor on the 2021 assessment for DOE. "Operating the grid with highly variable renewables is a massive problem." Still, he says, harnessing just 10% of the US's tidal energy resources could be an optimistic goal even as the technology matures.

Rising tides

The first commercially licensed US-based tidal turbine, Verdant Power's Roosevelt Island Tidal Energy (RITE) project, operated for nine months in 2020 and 2021. Three turbines, fixed to a platform on the river bottom, were installed at a depth of 10 m in the tidal strait of the East River, which connects Long Island Sound with Upper New York Bay. The turbines generated electricity when the water flow velocity exceeded 1.0 m/s, a total of 4.5 hours each day. They delivered more than 312 MWh to the electricity grid, enough to supply 40 homes, which on average use 900 kWh in a month. During the demonstration, Verdant Power honed its maintenance capabilities. "Showing they could extract and replace a turbine and nailing down what the costs would be for the operation was a big success," says Kilcher.

BEN LOEFFLER/ALASKA CENTER FOR ENERGY AND POWER



DRIFTWOOD AND PLANT MATTER can jam tidal turbine blades. Researchers at the Tanana River Hydrokinetic Test Site in Alaska are working on designs that minimize debris fouling.

European projects are more advanced because they attracted investment earlier. The first grid-connected tidal array was installed off Scotland's Shetland Islands in 2016 by Nova Innovation, which added a sixth turbine in January, making the array the largest in the world. Each turbine provides enough electricity for 72 homes and is cost-competitive with diesel generation, according to CEO Simon Forrest. In thousands of hours of underwater video footage, the company recorded fish feeding around the structures at slack tide, moving away when the current began to flow, and disappearing from the area when the turbine blades picked up speed. An independent Scotland environmental regulator determined that there were "no negative impacts," says Forrest. In Shetland, tidal energy is "still more expensive than solar or wind, but prices are coming down, and we believe that we'll be competitive with nuclear energy," he says.

Curran Crawford codirects the Pacific Regional Institute for Marine Energy Discovery (PRIMED) at the University of Victoria. In April the institute won Can\$2 million from British Columbia to continue work on supporting remote communities, many of which are First Nations, through marine renewables including tidal energy. PRIMED is now working to implement a tidal turbine testing facility in a channel along the British Columbia coast at the family-owned Blind Channel Resort, which welcomes the idea of demonstrating a tidal turbine. The tidal turbine will be integrated into the resort's microgrid along with batteries to reduce diesel usage. The researchers plan to deploy a resident turbine capable of producing approximately 50 kW at maximum current, and want to host other turbines over time to prove out operations in British Columbia waters. First Nation communities in similar settings are intrigued about switching to a climate-friendly energy source, but they are still wary of the technology and want to see demonstrated success, Crawford says.

Small fossil-fuel-dependent communities along rivers also benefit from turbine technologies and can provide lessons for tidal installations. In Alaska, turbines deployed in the Yukon River at the towns of Eagle and Ruby highlighted a problem in 2013. Their blades frequently got jammed with logs, detritus, silt, and other debris, says Ben Loeffler, who leads the hydrokinetic research group at the Alaska Center for Energy and Power. To help develop technologies that avoid such jamming, he and his colleagues work at a test site on the Tanana River, where facilities let researchers quantify debris loading and make changes to turbine designs. What they learn could benefit future offshore projects: South-central Alaska's Cook Inlet is home to one of the most powerful tidal energy sites in the US. "We're translating our river turbine experience and knowledge to the Cook Inlet now that it is being looked at seriously for tidal energy," says Loeffler.

Technical creativity

New approaches are improving turbine efficiency and reducing blade damage. Some lessons can be learned from waterwheels that have their support structures above water and can be accessed from the land or a boat, in contrast to the seafloor foundations that support bladed turbines. But there are drawbacks. "Waterwheels are simple and cost-effective, they don't have blades that could damage fish and marine mammals, but until you find a way to increase the flow, they're not efficient," says Gerald Muller, an associate professor of engineering and the environment at the UK's University of Southampton.

Some researchers in the tidal energy community believe waterwheels can be refined to provide the functionality of a turbine and may provide some advantages. Working with Muller as a scientific adviser, Halifax-based BigMoon Power designed a floating, flow-augmented waterwheel. Inclined plates create a low-pressure zone on the wheel's downflow side; the zone lowers the water level in the wheel's channel and increases flow velocity by 50%. Because the kinetic power of a fluid is a cubic function of the flow velocity, the available power is increased by a factor of 3.4. The floating platform is anchored with removable concrete blocks, keeping all mechanical and electrical parts above water for easy maintenance while avoiding the leakage and condensation that plague turbine electronics.

"In the offshore business, it's a rule that everything that can break will break. You want to keep it as simple as possible," says Muller.

BigMoon Power has secured a license to operate eight of its 500 kW tidal energy generators in the Bay of Fundy, which separates New Brunswick from Nova Scotia and has the highest tides in the world. The team is installing a second-generation version of their 12.6-m-diameter device that will connect to Nova Scotia's grid this summer.

The debris in Alaska's Yukon River led Loeffler's colleagues to create a design that could harvest energy from turbulent flow at sites they had previously considered untenable. The researchers developed a vertical oscillator system that draws energy from a pipe placed transverse to the currents. Although the device turned out not to be economical and was discontinued, other companies are developing similar systems for river and tidal currents. Another design deployed and tested by Oregon's BladeRunner Energy in the Tanana River keeps a rotor submerged at the end of a large flexible driveshaft. The rotor transmits torque while floating at the end of a cable in the water. That design is on course for a pilot project in an Alaskan community in 2025. "Their 5 kW system could power five homes in the summertime," says Loeffler.

Economic viability

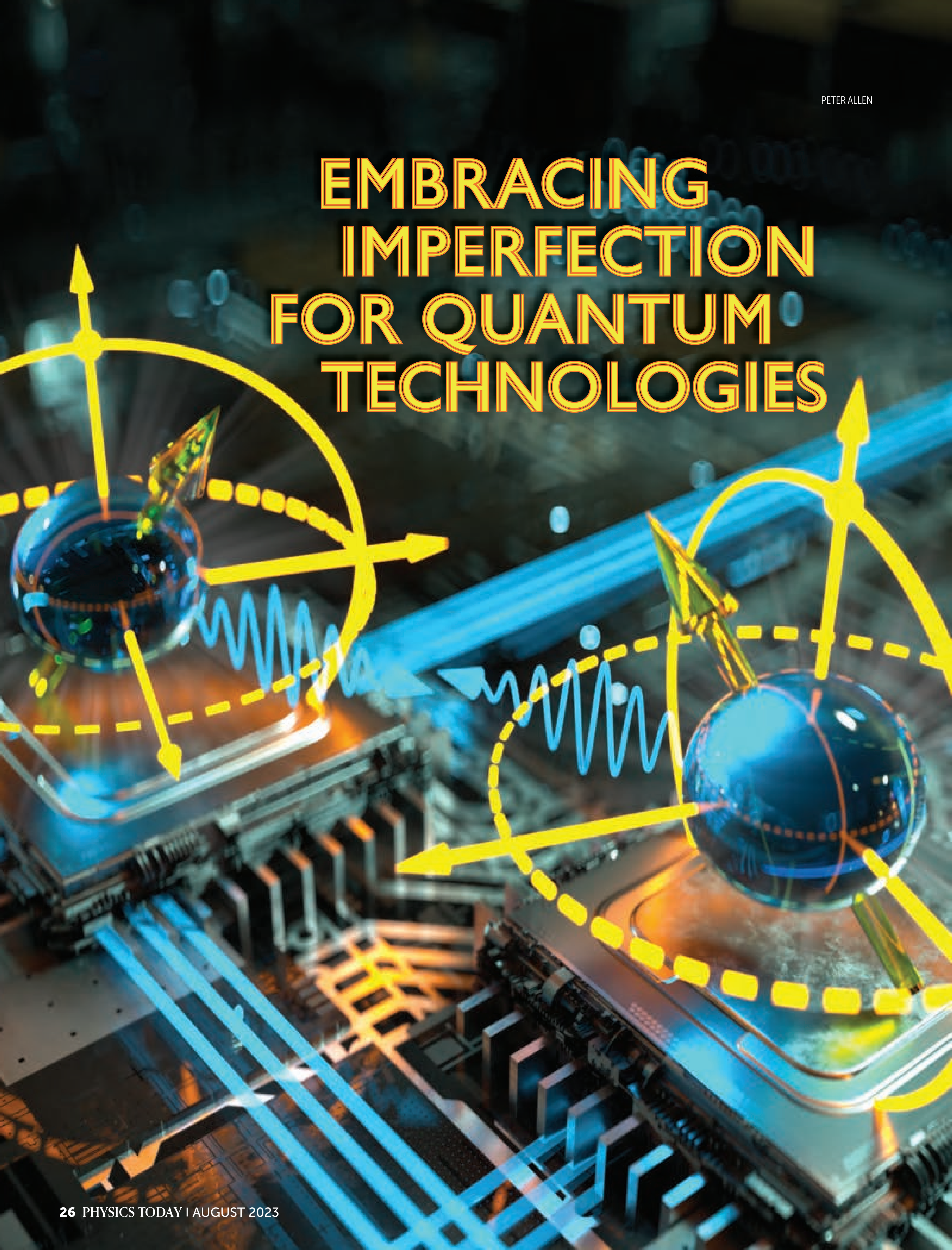
Efforts to deploy megawatt-scale tidal turbines "have been challenged by getting enough capital together," says Crawford. "If something goes wrong, you sink the company." In the near to medium term, he says, the most promising market segment is smaller turbines for decarbonizing remote communities. They are a necessary step to multi-megawatt-scale versions. "Sidestepping small device development introduces risk," says Crawford.

According to Kilcher, the energy sector considers a project that is robust and reliable for 20–30 years to be economically viable. "In wind turbines, this lifetime is the standard," he says, "and it's where we need to be for tidal turbines."

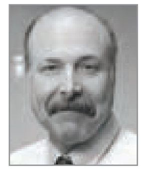
Rachel Berkowitz PT

PETER ALLEN

EMBRACING IMPERFECTION FOR QUANTUM TECHNOLOGIES



Christopher Anderson is an incoming assistant professor of materials science and engineering at the University of Illinois at Urbana-Champaign. **David Awschalom** is the Liew Family Professor of Molecular Engineering and Physics at the University of Chicago and a senior scientist at Argonne National Laboratory in Lemont, Illinois.



Christopher P. Anderson and David D. Awschalom

Solid-state spin qubits unlock applications in nanoscale quantum sensing and are at the forefront of creating distributed, long-distance entanglement that could enable a quantum internet.

You probably think something being defective makes it worse. Surprisingly, in the quantum world, defects in materials can be used as robust quantum bits, or qubits, that could fundamentally change the way we process, store, and distribute information. While undesirable in traditional semiconductor devices, lattice imperfections—including vacancies and impurities such as unwanted dopants—are now being harnessed for their unique quantum properties. In fact, they are leading candidates for developing potent quantum technologies, such as unlocking magnetic imaging at the nanoscale; for enabling a new internet of powerful quantum computers; and for creating unhackable communications secured by the fundamental laws of physics.

The irony that material imperfections are potentially ideal qubit candidates is not lost on the scientific community. For more than 50 years, the electronics industry has spent countless hours and resources trying to eliminate defects. They can affect the performance of everyday technologies like the processor chips in your computer and the integrated circuits making your car safe—potentially causing them to malfunction. Now those very same defects are being deliberately introduced into materials to create qubits with state-of-the-art quantum properties.

Defects come in a vast variety of forms.[1] The distinguishing features that can be harnessed for quantum technologies are the presence of unpaired electrons and multiple orbitals that those electrons can occupy. Electrons can undergo optical transitions between a defect's orbitals by absorbing and emitting single photons at specific wavelengths, which is why they are also sometimes called color centers. The emission can be measured using standard optical microscopy techniques—defects show up as luminescent dots in the crystal (see figures 1a and 1b).

Importantly, the electron undergoing optical transitions also possesses a magnetic moment whose spin-up and spin-down states act as a prototypical quantum two-level system—a qubit. The electron's spin state and orbital structure, and therefore its functionality, are defined by both the host crystal and the defect type (see figure 1c). Largely isolated and protected from their environment, the spin and orbital states are effectively embedded in a "semiconductor vacuum."

Defects as qubits gained prominence more than 20 years ago with the discovery of the many attractive

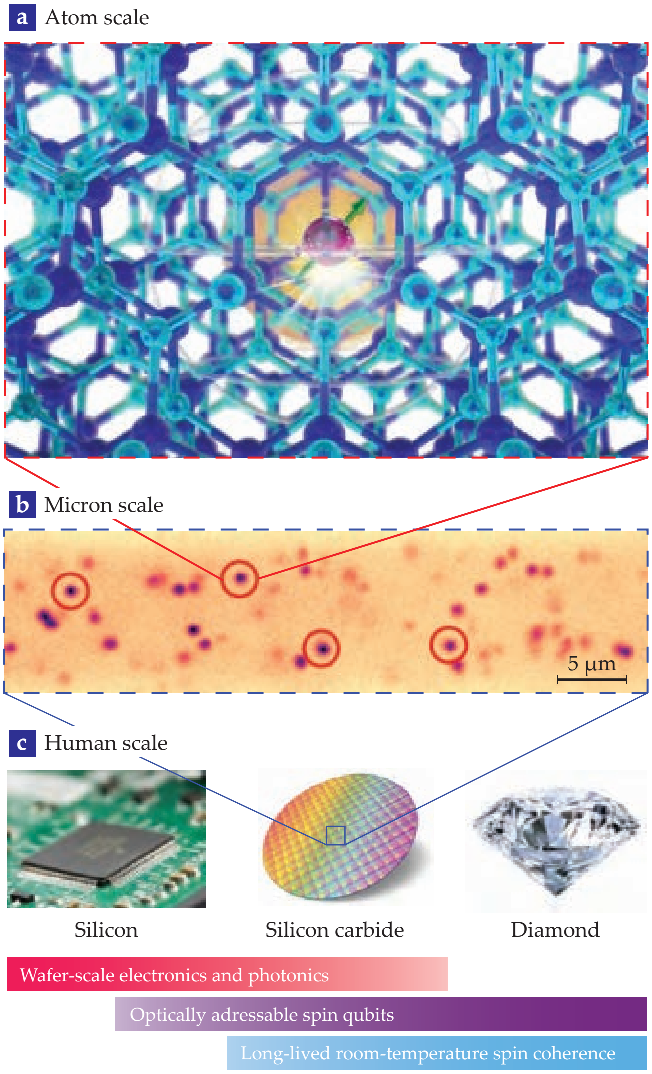


FIGURE 1. CRYSTAL DEFECTS AS QUANTUM BITS. (a) Schematic atom-scale representation of a defect spin qubit (glowing purple ball) in a crystal lattice. **(b)** In this micron-scale scanning photoluminescence measurement of isolated defects in silicon carbide, the circled dots are examples of confirmed optically addressable single qubits. (Adapted from K. Miao et al., *Sci. Adv.* **5**, eaay0527, 2019.) **(c)** Example host crystals for defect-based qubits are silicon, silicon carbide, and diamond. The colored bars indicate which systems display various desirable features.

features of the nitrogen–vacancy (NV) center in diamond, in which two adjacent carbon atoms are replaced by a nitrogen atom and a vacancy (see the article by Lilian Childress, Ronald Walsworth, and Mikhail Lukin, Physics Today, October 2014, page 38). NV centers in diamond can store quantum information for extended periods of time, and the quantum state can be read out optically at room temperature. Those properties make NV centers an ideal platform for quantum sensing of pressure, temperature, magnetic fields, and electric fields. They have also been used to close loopholes in Bell's inequalities[2] and to perform rudimentary quantum error correction for quantum computation.[3]

The defect trifecta

Defect qubits combine three attractive features: They're spin based, they're optically active, and they're solid state.

Why spins? In the context of quantum information technologies, a qubit's lifetime and coherence time—how long its entanglement and superposition last—are everything. They put a fundamental upper limit on the number of logical operations that can be performed before the state is lost and set the sensitivity with which a quantum system can detect its environment. By those benchmarks, electron spins can be extremely robust quantum objects, potentially having lifetimes of hours and coherence times of seconds at low temperature. Even in noisy room-temperature environments, the coherence times can be longer than a millisecond, which is sufficient for most quantum applications. Additionally, the spin's energy can be easily tuned with static magnetic fields and its quantum state controlled by oscillating fields—the same spin-resonance techniques utilized in MRI, for example. In many ways, electron spins are "textbook" qubits, and high-fidelity control with off-the-shelf RF electronics is straightforward.

Why light? The linking of the spin magnetic moment to optical transitions is the key feature that distinguishes quantum defects in solids from other spin-based quantum technologies (see the article by Lieven Vandersypen and Mark Eriksson, Physics Today, August 2019, page 38, for comparison). That critical spin–photon interface, described in box 1, has important implications.

For starters, light allows for highly nonequilibrium, nonthermal spin polarizations. In other words, it enables efficient initialization of the spin into a particular quantum state even in "hot" environments. Spin transition frequencies are commonly in the gigahertz range, or millikelvin in units of temperature. So at room temperature, spins in thermal equilibrium are not polarized into a particular state. They are instead in a probabilistic, nearly equal mixture of spin-up and spin-down states. Even at the liquid-helium temperature of 4 K, spins are usually polarized only weakly, by less than 5%. Thermally initializing a spin qubit would require ultrahigh magnetic fields, to provide a large energy difference, or ultralow temperatures—constraints that greatly limit practical usage.

Light provides an alternative solution for qubit initialization. The high energy (hundreds of terahertz) of the optical transitions means the excited states are not thermally occupied, and laser-driven, spin-dependent optical processes can produce ground-state spin polarizations approaching 100% even at room temperature. The coupling to light is what makes quantum operation possible at elevated temperatures, since initializing the qubit into a known state is the essential first step in all quantum protocols.

Optical transitions also enable the accurate, deterministic readout of the quantum state. Because the magnetic dipole moment of a single spin is small, it couples only weakly to its environment, making spin a difficult quantity to measure directly—doing so requires extremely low temperatures, high fields, and complicated device integration. On the other hand, using suitable optical transitions can make readout of single spins easy; it can even be a demonstration for teaching purposes in introductory undergraduate physics labs. That ease (not guaranteed for all defects) arises from the specific quantum mechanical selection rules for the allowed transitions of the spin–photon interface. Importantly, the interface also un-

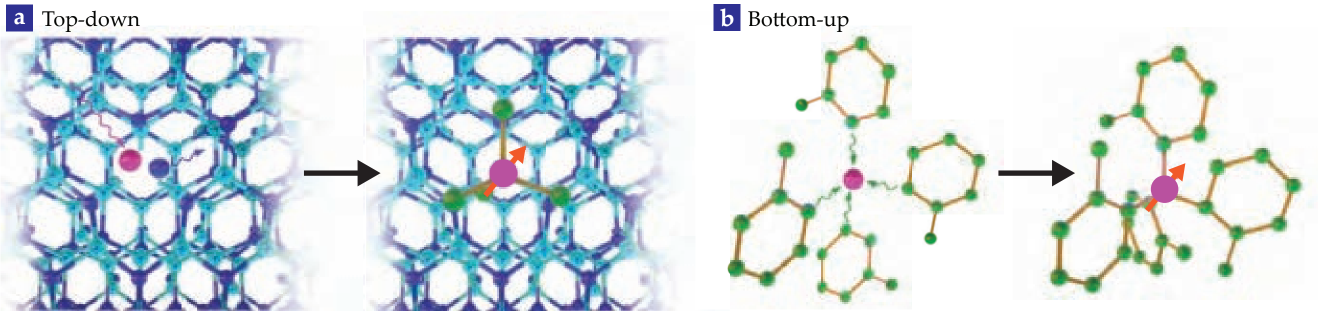


FIGURE 2. CREATING DEFECT SPIN QUBITS. (a) Most spin systems today are created with a "top-down" approach where impurities or vacancies are introduced into an existing structure. Here, an impurity (magenta) is introduced into a lattice, displacing an atom (blue) from a lattice site. The impurity and vacancy together can function as a qubit. **(b)** A "bottom-up" approach instead builds the entire structure atom by atom. Schematically, a central ion (purple) can be coordinated with specific chemical ligands (green) to create a structure with the desired behavior.

locks the ability to entangle single electron spins with single photons (see box 1).

Finally, the weak, short-range dipolar interaction between spins makes coupling two spins together difficult. Thanks to the spin–photon interface, light provides an alternative for the long-range "wiring" needed for spin qubits to communicate with each other and with other quantum systems. Because the real power of quantum technologies comes from connecting and entangling many qubits, such long-range interactions are essential. Photons are also the ideal transmitters of quantum information: They travel at the speed of light, can be guided in low-loss optical fibers, and are noise-free at room temperature. Excitingly, such light-mediated entanglement of distant defect-based qubits to form rudimentary quantum networks has

Box 1. The spin–photon interface

While the spin-up and spin-down ground states of a solid-state defect form the basis for a qubit, it is the optical transitions between those ground states and higher-energy orbitals that enable qubit preparation, readout, and coupling between qubits.

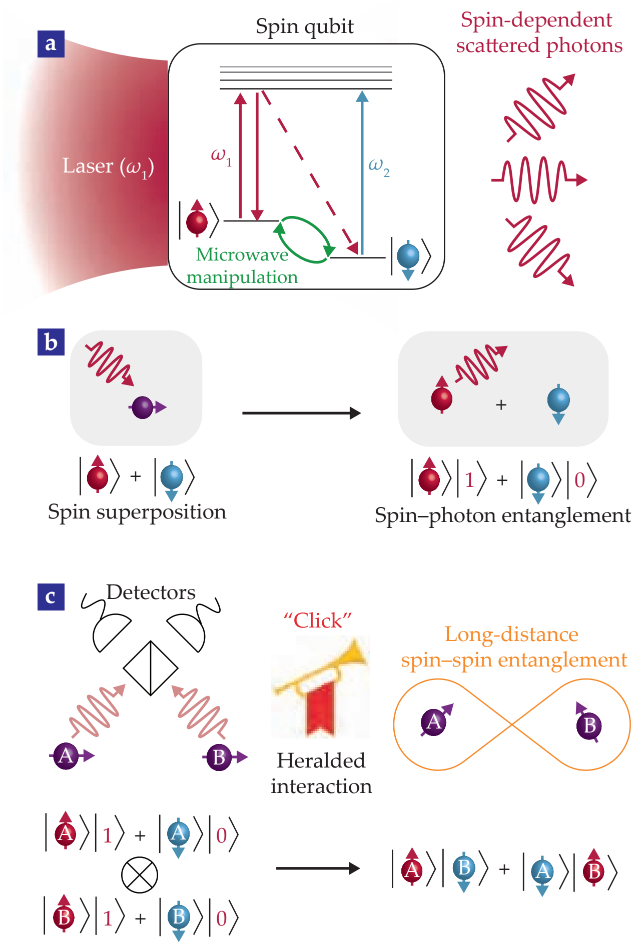
Spin-selective optical transitions between the ground state and excited states allow for spin-dependent photoluminescence and state readout, as sketched in panel a. Here, the frequency ω_1 that will optically excite a spin-up electron (red) will not excite a spin-down electron (blue). The excited defect will emit light when it relaxes, so the detection of scattered photons indicates that the spin is up.

Spin-state preparation is achieved by pumping the defect with a laser tuned to frequency ω_1. An optically excited spin-up qubit has a finite probability of flipping (dotted red line) as it returns to the ground state. Once the qubit is spin-down, the laser is no longer resonant, and the state is stationary. Therefore, after sufficient pumping, the qubit has a nearly 100% probability of being spin-down. Microwave signals (green) tuned to the energy separation between spin states can manipulate the qubit state using common electron spin-resonance techniques.

In this example, the spin dependence comes from frequency selectivity, but more generally a spin–photon interface can be selective based on any property of light, including polarization.

Panel b illustrates how spin-selective optical absorption can transform a spin superposition state (purple, left) into an entangled state between the spin and the presence (1) or absence (0) of a single photon. Because only the spin-up state couples to incident light at ω_1, a photon is created only when the spin is up.

Photons emitted by the spins can mediate long-range interactions, as sketched in panel c. Consider two spatially separated spin qubits that, using the scheme in panel b, are entangled with their emitted photons and have the same frequency. Interference at a 50:50 beamsplitter and subsequent detection ("click") of a single photon will announce, or herald, the creation of entanglement between the distant spins. As a result of the measurement, the initial pair of entangled spin–photon states is transformed into an entangled state of just the two spins.



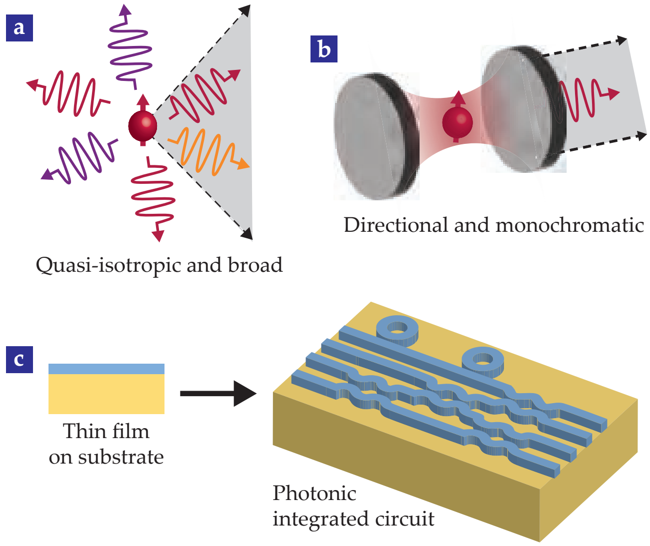


FIGURE 3. PHOTONICALLY ENHANCED QUBITS. (a) In free space, defect-based single-photon sources tend to emit in many directions and have a wide spread of emission frequencies (different colors). Only a small fraction of the emitted light is collected (gray cone). **(b)** The emission behavior can be enhanced by integrating the source into an optical cavity, schematically displayed as two mirrors. Photon emission gets funneled into the cavity mode, dramatically decreasing the frequency spread and increasing the useful light collected. **(c)** Integrated photonic devices can be created by starting with a high-quality thin film of the desired qubit-hosting material (blue) bonded to a substrate (yellow) with lower index of refraction. The top layer can be patterned into photonic integrated circuits that contain the necessary components—cavities, beamsplitters, phase shifters, and so on—to perform quantum optics experiments on a chip.

recently been demonstrated.[4] In a sense, defects provide built-in strong coupling between the spin and light degrees of freedom and thus enable efficient translation of quantum information across disparate frequencies.

Why solid state? Hosting quantum states in a solid material gives unique advantages. First, the power and scalability of semiconductor technology can be tapped to naturally embed qubits in devices fabricated from the host material. For example, micro- and nanofabricated devices that control photons, phonons, or charges can be used to manipulate, protect, or couple qubits. Second, scientists can create and manipulate isolated qubits without the ultrahigh vacuum or ultrastable lasers typically used in trapped-atom and trapped-ion experiments. As a result, defect-based quantum systems and devices can be small, portable, and usable in ambient environments—desirable traits for practical quantum devices for the real world.

As an added benefit of working with solid-state systems, the electron spin states of defects can couple to nearby lattice atoms that, because of their isotope, have a net nuclear magnetic moment. Such nuclear spins in solids are perhaps one of the most robust quantum memories known to science, with coherence times of minutes and lifetimes that can exceed days.[5] Local multiqubit entangled registers of single electrons and nuclei have been created in a multitude of solid-state defect-spin platforms.

The combination of scalable fabricated devices, long-lived spin states, and light-based initialization, readout, and mediation of interactions is what sets defect qubits apart. For example, the ability to engineer systems in an integrated, solid-state platform allows for unprecedented control of the magnetic, electrical, and photonic states of the qubit needed to optimize performance. Just like how the modern information age uses magnetic states (for longevity) in hard drives to store data, light (for speed and bandwidth) to communicate, and semiconductor devices (for scalability and size) to compute, defects in solids similarly leverage the same advantages for quantum technology.

Beyond the NV center

Over the past two decades, the field of quantum information has expanded greatly, and scientists now have a broad range of quantum systems at their disposal—they can choose which qubit to employ based on their needs and the specific quantum application. Atomic physicists, for example, can select from several different atoms and ions from the periodic table. Meanwhile, superconducting systems have evolved beyond the first charge qubits and now offer an impressive variety of artificial two-level systems, each with unique protections from noise. Such advances in quantum science have arisen from a pioneering spirit to discover new platforms and to understand the underlying fundamental physics to mitigate or even circumvent shortcomings in performance. In the same way, the field of defect-based spin qubits has expanded beyond the originally discovered NV center in diamond, as illustrated in figure 1c.

In the search for new and better qubits, a physics-based understanding of the desired properties of solid-state defects has served as a guide for more than 20 years.[6,7] The widening of the field's scientific scope beyond the NV center resulted in the discovery of optically addressable defects in silicon carbide[8] (see Physics Today, January 2012, page 10). Those defects act almost in direct analogy to NV centers in diamond. In particular, they have the features previously thought unique to the NV center: room-temperature initialization and readout of single electron spin states with long coherences. In addition, SiC has many of the desirable properties of diamond, including hardness, stiffness, optical clarity, high thermal conductivity, and a high index of refraction. As a result, SiC is used as a substitute for diamond both in jewelry and in industrial applications such as abrasives.

Where SiC really shines, however, is as a technologically mature, wafer-scale semiconductor; it is used for components in electric cars, 5G technologies, and LED light bulbs. Billion-dollar SiC fabrication facilities have opened in the past year that can produce wafer-scale, quantum-grade single-crystal materials. Besides clear cost and scalability advantages over diamond, SiC is also a much easier material from which to fabricate useful quantum devices. In fact, SiC is compatible with the same CMOS fabrication techniques that are used by the semiconductor industry to make today's microchips. Quantum states in SiC can be readily integrated into the mechanical resonators found in a cell phone, into electrical devices to read out their quantum state in new ways, and into low-loss photonic circuits to create efficient entanglement, among other potential applications.

All the key quantum functionalities have been demonstrated in SiC over the past 10 years or so: single-qubit creation, high-fidelity control, nuclear-spin quantum memories, device integration, single-shot quantum readout, and record-long coherence times.[9] In many ways, SiC has caught up with or exceeded

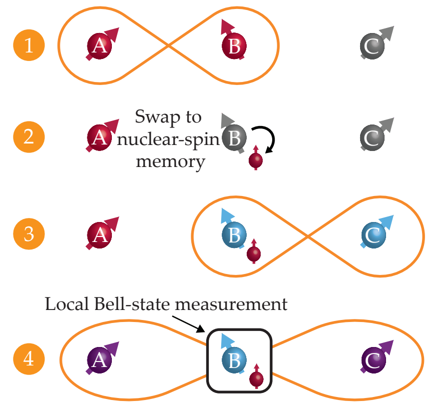
Box 2. Enabling long-range quantum networks with defects

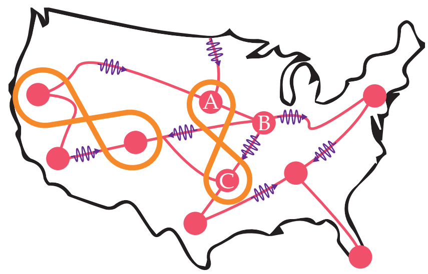
Any future long-range quantum network will require quantum repeaters to prevent information loss. Solid-state defects can implement such a quantum repeater via a four-step protocol:

1. Defects A and B are entangled (denoted by the orange lemniscate) using the photon-brokered interaction described in box 1.
2. Defect B's half of the entangled state is swapped into a long-lived nuclear-spin quantum memory.
3. Defects B and C are entangled. Even if the interaction fails, the entanglement from step 1 is preserved by the quantum memory and the entanglement can be attempted again.
4. A local two-qubit measurement (called a Bell measurement) of the electron and nuclear spins at B "fuses" the separately entangled pairs (red and blue) and creates an overall entanglement between defects A and C (purple).

The quantum repeater node at B obviates the need for direct, coherent transmission between A and C; it instead breaks up the quantum link into independent segments and allows entanglement to be created even though A and C never directly interact.

Such defect-based repeaters could form the backbone of long-distance quantum networks, thereby enabling a distributed internet of entangled quantum devices. Quantum nodes could be linked by light sent over a fiber-optic network, sketched here as red circles and lines. The defects described above, for example, could be located in the nodes labeled A, B, and C.





the performance of NV centers in diamond—exemplifying the power of searching for qubits in new materials.

Scientists have since expanded their search to include new types of defects in diamond, in silicon,[10] and even in two-dimensional materials such as hexagonal boron nitride.[1] The unique properties of these new host materials and defects have driven many of the major advances in the field. Understanding the vast interplay between the desired application, host crystal, and defect structure underpins a recent "blueprint" for quantum materials science with defects[11] and a new "periodic table for quantum coherence."[12] For example, calcium tungstenate ($CaWO_4$) was predicted and recently shown to have record-long spin coherence times for a naturally occurring crystal. It's also part of a broad trend of repurposing traditional laser crystals, including rare-earth-doped oxide materials, for quantum applications. In fact, many of the features that make a good gemstone—including a high index of refraction, optical clarity, and hardness—are exactly the features one wants in a host of solid-state defects.

Most solid-state defect qubits are made with a "top-down" approach, in which a crystal host is implanted or irradiated to create impurities or vacancies, as shown in figure 2a. Those processes make it difficult to place qubits in desired locations and may cause unwanted damage to the crystal, which results in reduced coherence. Tackling those challenges is an ongoing area of research. Furthermore, top-down approaches are subject to fundamental design limitations set by thermodynamics, crystal symmetries, commercial availability of the crystals, and more.

One emerging alternative is to build optically active solid-state spin systems from the "bottom up," as shown in figure 2b. Instead of removing atoms or adding single impurities to a host, one can use the power of synthetic chemistry to assemble a quantum system atom by atom. For example, researchers have designed transition-metal molecular complexes with optical transitions, and by changing the ligands in the complex they tuned the spin and optical properties.[13] A bottom-up strategy has even been used to tailor the level structure in molecular complexes to increase quantum coherence times. Finally, molecular complexes displaying coherent and narrow optical emission have also been demonstrated. Although relatively new, bottom-up approaches could offer unparalleled flexibility and potentially allow for single-molecule quantum devices designed with a particular quantum application in mind.

Photonic integration

Coupling to light is a distinguishing feature of solid-state defect-based quantum systems. Defects in solids, however, are not ideal single-photon sources. Photon emission is commonly disturbed by phonons, can be plagued by competing nonradiative relaxation pathways, or may be affected by nonidealities in the energy-level structure outlined in box 1. In some cases, the optical lifetime—the delay before a photon is emitted—can be so long that the qubit shines too dimly to be practically measured. In addition, the qubit is much like a light bulb: It radiates photons in all directions. If those photons contain information one wants to sense or they are mediating an interaction with another qubit, one does not want them to emit randomly (see figure 3a). Creating devices that enhance or guide the optical emission from defect spin qubits is therefore critical.

Luckily, there is a solution to all those problems at once. Edward Purcell noted in the 1940s that the spontaneous emission of a two-level system can be modified by embedding it inside a resonant circuit or cavity (see figure 3b). The resonant structure modifies the possible states that the system can emit light into, and it causes preferential channeling of photons into a single spatial and spectral mode defined by the cavity. This "Purcell effect" can also increase the emission rate, thus creating a brighter and purer source of single photons. Purcell noted that for the effect to be prominent, the cavity must be both small and low loss. The natural integration of defect-based qubits into a solid-state material is thus a major benefit, since the material's high index of refraction can confine light at the nanoscale.

High-quality nanoscale optical resonators with embedded defects can be created using the same technique that dominates the integrated-photonics industry: a thin film of material, usually silicon, on top of a substrate, often silicon dioxide, that has a lower index of refraction. For example, thin films of both SiC and diamond on oxide have recently been developed. Through subsequent simple vertical etching, one can make waveguides—essentially wires for light—that are surrounded on all sides by lower-index media, as shown in figure 3c.

Nanofabricated photonic devices have other advantages as well. For example, detectors, laser sources, and even linear optical elements such as beamsplitters can all be fabricated on a single chip. By combining this capability with nonlinear optical frequency conversion and electro-optic tunability of some materials, one can harness the full power of integrated photonics and achieve powerful quantum functionality. Nonclassical squeezed states of light could be created in the same resonator that hosts single-photon emitters, for instance. Defects can also potentially be sources of highly entangled states of light called cluster states, which can form the basis of measurement-based quantum computation, in which quantum algorithms are implemented via a series of single qubit measurements.

Mitigating noise

While some defects make robust qubits, not all imperfections are desirable for quantum applications. Unwanted lattice defects, extra charges, and uncontrolled spin magnetic moments can introduce noise that disturbs the quantum state. The spin ground state of a defect qubit is predominantly sensitive to nearby nuclear and electron spins in the host crystal, which cause fluctuating magnetic fields that limit qubit coherence. On the other hand, the coherence of single-photon absorption and emission, which sets the fidelity of state preparation and readout and of long-distance entanglement generation between spins (see box 1), is degraded by electrical noise, which arises naturally from the movement and fluctuations of electrons and holes in the solid. Electrical noise is especially challenging in fabricated devices with nearby surfaces that can trap charges.

For magnetic noise, purifying the host material so that it contains only specific isotopes with no nuclear magnetic moment has greatly increased coherence, which improves quantum sensing protocols and boosts quantum memory times. Spin qubits can also be made insensitive to magnetic noise by engineering energy levels so that they do not shift with external perturbations. In addition, one can protect coherence by using dynamic quantum control to rapidly flip the quantum state and produce a noise-cancellation effect known as a spin echo. A combination of such techniques recently achieved coherence times exceeding five seconds in SiC—the longest ever demonstrated for an electron spin in a solid, sufficient for nearly all desired quantum applications.[9]

Because light-based interactions are key to applications of defect spin qubits, combating electrical disturbances that degrade photon coherence is a priority. For diamond, there has been a pivot away from the established NV center toward vacancy centers involving silicon, germanium, and tin. The main advantage to using those group IV–based defects is inversion symmetry: The system is nonpolar and insensitive to electric fields to first order. That simple change has yielded dramatic improvements—robust single-photon emitters can be integrated into nanophotonic cavities and display highly coherent emission of light coupled strongly to the optical mode of the cavity.[14] For SiC, one can use doping to create a simple electrical diode, which has a depletion region that is completely devoid of free carriers and thus of fluctuating charges. The depletion region therefore eliminates electrical noise, resulting in near-perfect, highly tunable quantum emission.[15] Such doping control is simply not feasible in diamond, which highlights the unique advantages new materials can provide.

Those examples illustrate the two strategies to combat noise on a qubit. One is to eliminate the source of the noise. That can sometimes be achieved with careful materials engineering, but it has stringent fabrication requirements. The other is to create a qubit that is insensitive to the noise source. In doing so, however, the qubit also becomes insensitive to many tuning knobs that are needed to control the system. The most appealing approach will likely be a combination of the two so that the qubit operates in a "Goldilocks zone" of sensitivity and controllability.

Maturing technologies

The past decade's advances in quantum information science and engineering with solid-state defects have been largely the result of a shift in mindset—toward exploring new qubit candidates in new materials and expanding the range of applications. In the future, a diversity of approaches is likely, which reflects the varied applications that defects can tackle. The development of new platforms requires a multidisciplinary approach that involves materials science, atomic physics, condensed-matter physics, electrical engineering, and advances in nanofabrication. In fact, linking with the very semiconductor industries that have worked to eliminate defects may be the key to scaling defect-based quantum systems.

Even as they have been the subject of much ongoing research, defect-based quantum states have started to enter the mainstream for core quantum applications: sensing, computing, and communications. Much of the leading technical work was demonstrated using the NV center in diamond, from small-scale error correction[3] to a lab-scale three-node quantum network that can teleport quantum states.[4] NV centers have been commercialized into scanning-probe nanoscale magnetometers, which have been used to measure fundamental condensed-matter phenomena.[16] They have also found use as biosensors[17] (see PHYSICS TODAY, August 2011, page 17) and as room-temperature vector magnetometers for navigation and geoscience.

Although they can find use in all the major areas of quantum science, solid-state defects really shine in two critical applications: room-temperature quantum sensing of magnetic fields and long-distance quantum communications and networking. For quantum sensing, major efforts focus on optimizing near-surface qubits, which can sense the environment more readily than deeper qubits but are subject to noisy interfaces. For quantum networking, the main hurdle is efficient mediation of long-distance entanglement between qubits, which is limited both by the quality of optical devices and by the noisy photon emission that results from integration into those devices.

A tantalizing goal on the horizon is a quantum internet,[18] which, in analogy to the classical internet we know today, would link quantum devices and distribute quantum information. It would, for example, enable more powerful, modular quantum computers, distributed quantum sensing of gravity, and quantum key distribution over global scales. (For more on quantum key distribution, see the article by Marcos Curty, Koji Azuma, and Hoi-Kwong Lo, Physics Today, March 2021, page 36.) The 2022 Nobel Prize in Physics was awarded for work with spatially separated entangled photons (see Physics Today, December 2022, page 14). Creating such entanglement at the metropolitan scale or larger, however, will require quantum repeaters to mitigate photon loss. These repeaters each combine a spin–photon interface with a long-lived quantum memory (such as a nuclear spin) to break up a quantum channel into multiple shorter links. By buffering signals using the quantum memories, entanglement swapping efficiently links the end nodes, as described in box 2.

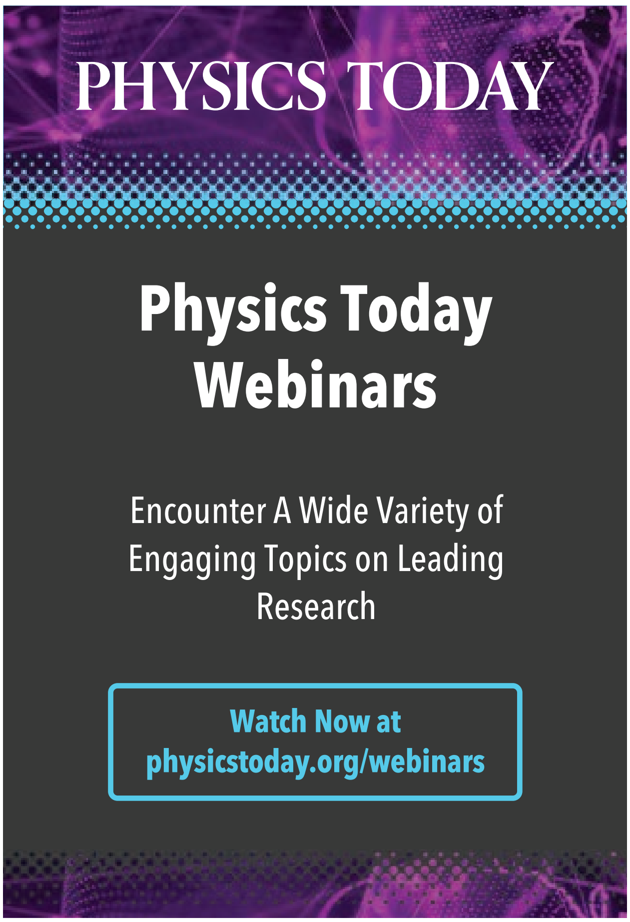
Defects in solids provide all the necessary components to implement quantum repeaters and create a scalable quantum network backbone: Photons are the natural choice for quantum communications, spins are robust memories, and semiconductors are scalable. Such a backbone is the focus of recent large industrial efforts, quantum startups, and national and international collaborations of government labs and universities. In the future, we envision rack-mounted cryostats stashed in closets across the country, with defect-based devices interfaced with fiber-optic cable to route quantum entanglement for sharing secure cryptographic keys, distributing quantum computation, and making more powerful sensors.

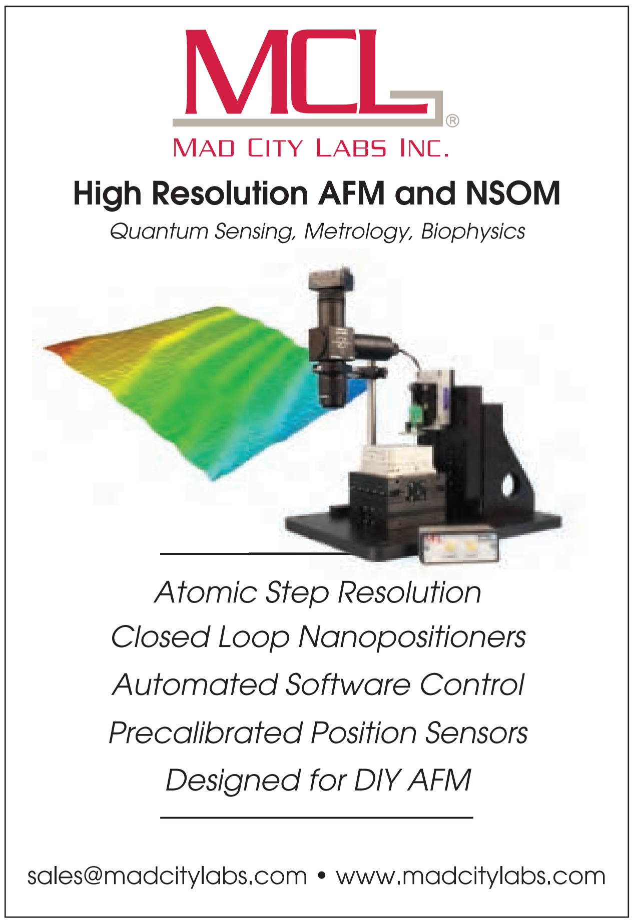
The authors' work on quantum science and engineering is supported by the US Air Force Office of Scientific Research, the US Department of Energy (Q-NEXT), and NSF. The research was also supported by the Intelligence Community Postdoctoral Research Fellowship Program at Stanford University administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the US Department of Energy and the Office of the Director of National Intelligence.

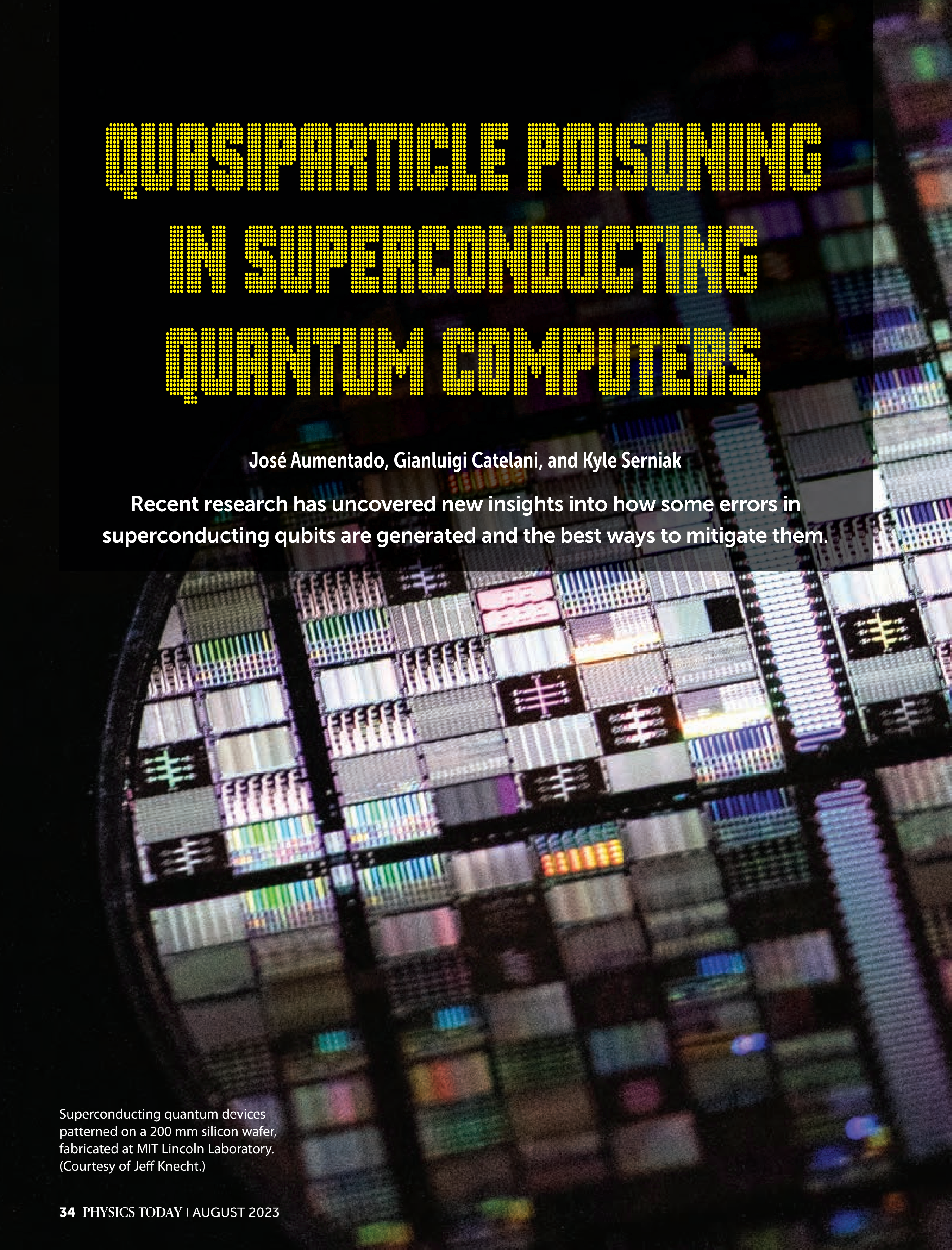
REFERENCES

1. M. Atatüre et al., *Nat. Rev. Mater.* **3**, 38 (2018).
2. B. Hensen et al., *Nature* **526**, 682 (2015).
3. M. Abobeih et al., *Nature* **606**, 884 (2022).
4. M. Pompili et al., *Science* **372**, 259 (2021).
5. K. Saeedi et al., *Science* **342**, 830 (2013).
6. D. DiVincenzo, *Fortschr. Phys.* **48,** 771 (2000).
7. J. Weber et al., *Proc. Natl. Acad. Sci. USA* **107**, 8513 (2010).
8. W. Koehl et al., *Nature* **479**, 84 (2011).
9. C. Anderson et al., *Sci. Adv.* **8**, eabm5912 (2022).
10. D. Higginbottom et al., *Nature* **607**, 266 (2022).
11. G. Wolfowicz et al., *Nat. Rev. Mater.* **6**, 906 (2021).
12. S. Kanai et al., *Proc. Natl. Acad. Sci. USA* **119**, e2121808119 (2022).
13. S. Bayliss et al., *Science* **370**, 1309 (2020).
14. A. Sipahigil et al., *Science* **354**, 847 (2016).
15. C. Anderson et al., *Science* **366**, 1225 (2019).
16. L. Thiel et al., *Science* **364**, 973 (2019).
17. N. Aslam et al., *Nat. Rev. Phys.* **5**, 157 (2023).
18. H. Kimble, *Nature* **453**, 1023 (2008).

PT







QUASIPARTICLE POISONING IN SUPERCONDUCTING QUANTUM COMPUTERS

José Aumentado, Gianluigi Catelani, and Kyle Serniak

Recent research has uncovered new insights into how some errors in superconducting qubits are generated and the best ways to mitigate them.

Superconducting quantum devices patterned on a 200 mm silicon wafer, fabricated at MIT Lincoln Laboratory. (Courtesy of Jeff Knecht.)

José Aumentado is a staff scientist in the Advanced Microwave Photonics Group at NIST in Boulder, Colorado, and is a senior fellow at Quantum Circuits Inc in New Haven, Connecticut. **Gianluigi Catelani** is a research scientist in the Peter Grünberg Institute at the Jülich Research Center in Germany and lead researcher in the Quantum Research Center at the Technology Innovation Institute in Abu Dhabi, United Arab Emirates. **Kyle Serniak** is a technical staff scientist at MIT Lincoln Laboratory in Lexington, Massachusetts, and at the MIT Research Laboratory of Electronics in Cambridge.

Although quantum computing is still in its infancy relative to the "classical" computing technology that we've come to know, love, and rely on, rapid advances over the past decade have taken it from the realm of science fiction to a probable reality of the not-so-distant future. Instead of manipulating bits of information by operating millions of transistors, a quantum computer relies on the precise control of many quantum subsystems—individual quantum bits, or qubits—along with an accurate readout of their quantum states. Many promising physical qubit platforms, such as trapped ions, neutral atoms, and solid-state defects (see the article by Christopher Anderson and David Awschalom on page 26), are based on building blocks that are typically thought of as archetypes of quantum behavior.

One of the leading candidate platforms for a useful quantum processor, however, is constructed from components that don't evoke a picture of tiny, microscopic particles with exotic properties. Instead, it consists of superconducting wires, capacitors, and inductors patterned on chips akin to existing semiconductor technologies. Those electronic circuits, which make up the superconducting qubit platform, embody many of the desirable properties of their atomic counterparts and have become the focus of several high-profile quantum computing efforts—led by both large companies, such as IBM, Google, and Alibaba, and startups, including Rigetti Computing, IQM, Alice & Bob, Oxford Quantum Circuits, and Quantware.[1] Those companies are leveraging modern clean-room fabrication tools to more easily engineer complex circuits with fast control.

In developing any quantum computing platform, a fundamental challenge arises from the tension between preserving quantum information and manipulating it: The former requires that qubits be isolated from their environment, while the latter demands that they have precise interactions with it. In fact, the key metrics for any platform can be summarized by the probability that an error will occur during a calculation and the time it will take to complete that calculation.

Currently, researchers looking at superconducting qubits are focusing on the error probability, which can be thought of as the ratio of how fast the qubit can be controlled to the rate at which it loses information to its environment. Of the primary mechanisms that are currently limiting superconducting qubit performance, one of the most intriguing and difficult to control is quasiparticle poisoning—the presence of charge carriers that do not participate in the superconducting condensate.

Superconducting qubits in a nutshell

Quantum effects are often weak and hard to observe in objects visible to the human eye. (For example, see Physics Today, July 2023, page 16.) So how is it that superconducting devices that are constructed from such circuit elements as inductors and capacitors and contain on the order of 10^{15} atoms behave quantum mechanically? As first shown by John Martinis, Michel Devoret, and John Clarke in 1987, a macroscopic degree of freedom can exhibit quantum behavior provided that energy dissipation is negligible and that the temperature of the system is low.

Thus the first ingredient to build a quantum circuit is to avoid energy dissipation, which leads to information loss. That's why circuit components are fabricated with superconducting materials. They can carry direct current without any resistance because the relevant charge carriers—electrons and holes near the Fermi energy—partner into Cooper pairs and condense into a macroscopic coherent state, as explained by the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity (see the article by Warren Pickett and Mikhail Eremets, Physics Today, May 2019, page 52). The condensate can be described by a complex-valued order parameter, the phase of which is critical to describe the physics of superconducting qubits.

Dissipationless transport is possible not only within bulk superconductors but also between two connected superconductors separated by what's called a weak link. The most widely used type of weak link is a tunnel barrier—a thin oxide layer separating two superconducting electrodes to form a Josephson junction. Importantly, Josephson junctions behave as nonlinear inductors: They lie at the heart of superconducting qubits, and the difference in the phase of the order parameter between the superconductors they connect is exactly the macroscopic degree of freedom that was shown to exhibit quantum behavior. In practice, aluminum is the superconductor of choice for Josephson junctions because it's compatible with relatively standard nanofabrication techniques and has a self-limiting few-nanometers-thick oxide at its surface, which is used for the junction barrier.

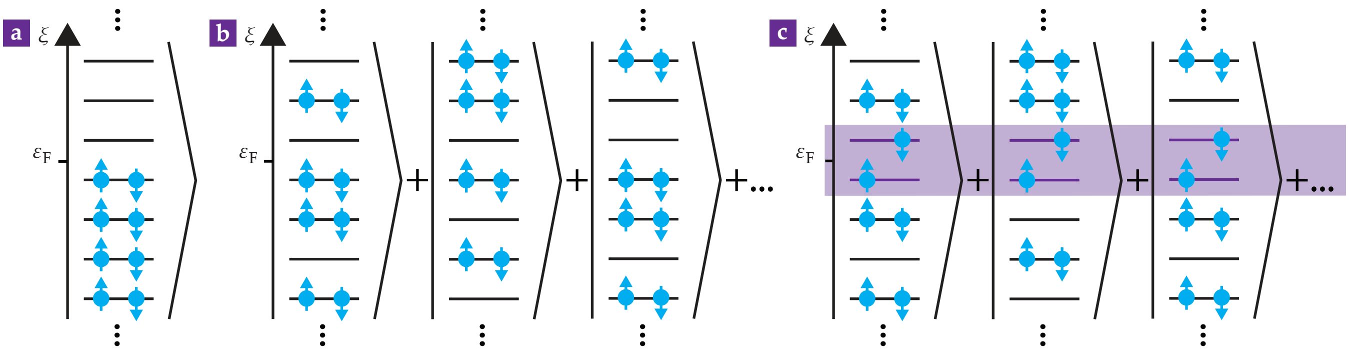


FIGURE 1. QUASIPARTICLE EXCITATIONS in superconductors. **(a)** In the ground state of a normal metal, spin-degenerate electrons (blue) occupy states with energy ξ up to the Fermi level ε_F. **(b)** The ground state of a Bardeen-Cooper-Schrieffer superconductor consists of a coherent superposition of all possible configurations of states, which have pair-correlated electron occupation in an energy window $\pm\Delta$ around the Fermi level. For simplicity, panels a and b neglect degeneracy or correlation in the momentum of the electrons. **(c)** When a phonon or photon with energy greater than 2Δ couples to the superconductor, the generated pair of quasiparticles poisons the superconductor: The two states the quasiparticles occupy (purple region) are fixed and don't participate in the coherent superposition of the superconducting condensate. (Adapted from ref. 9.)

The design flexibility of superconducting circuits originates from the many possible ways of combining the three basic circuit elements—capacitors, linear inductors, and nonlinear inductors (Josephson junctions), which all have parameters that can be tuned over a wide range. Is there a price to pay for such flexibility? Depending on how the components are arranged, quantum information can be encoded into the charge or the phase difference between superconducting condensates or as a combination of the two. The encoding methods hint at what can go wrong: The charge, the phase, or even the superconducting condensate itself can be disrupted.

Broadly speaking, the environmental effects acting on the charge or phase are known as charge noise and flux noise, respectively. They arise from materials defects and imperfections on the surface of the superconductor, at the interface with the substrate, in the oxide forming the Josephson junction, and in the substrate itself. At the microscopic scale, the sources of charge noise and flux noise arise from random changes in the configurations of charges and electron or nuclear spins.[2]

Another decoherence mechanism affecting charge and phase arises from the interaction of the superconductor with the electromagnetic environment: Like any other resonant electric circuit, a superconducting qubit can lose energy by emitting a photon. That's easy to visualize for the simplest superconducting qubit, called a transmon. Consisting of a Josephson junction in parallel with a capacitor, a transmon can be thought of as a nonlinear dipole antenna, which absorbs and emits photons at some characteristic frequency.

In contrast to the decoherence mechanism described above, the superconducting condensate can be directly disturbed by the environment via the breaking of Cooper pairs, a process that generates quasiparticle excitations in the superconductor itself. Cooper pairs comprise two electrons with opposite spin and momentum, and superconductivity results from the coherent superposition of the underlying many-body momentum states, which are either pair-occupied (electrons) or pair-unoccupied (holes), as illustrated in figure 1.

Picturing quasiparticles as broken Cooper pairs gives an idea of what they actually are. In a normal metal, electrons occupy various energy levels in a so-called Fermi sea, and when an electron is removed, what's left is a hole excitation. When removing an electron that was part of a Cooper pair, what's left is a coherent superposition of an electron and a hole, known as a Bogoliubov quasiparticle.

Whereas any small amount of energy is sufficient to generate an electron and a hole in a normal metal, it takes a finite energy, denoted as 2Δ, to break a Cooper pair. That energy, known as the superconducting gap, is proportional to the critical temperature T_c at which the superconductivity disappears: $\Delta \simeq 1.76\, k_B T_c$ for well-behaved BCS superconductors, such as aluminum. Because of the energy gap, at low temperature the thermally activated number of quasiparticles, which can be quantified as the fraction x_{QP} of broken Cooper pairs, should be exponentially small, $x_{QP} \sim \exp(-\Delta/k_B T)$. For aluminum at about 20 mK—the temperature at which aluminum-based superconducting qubits are typically operated—x_{QP} is expected to be about 10^{-46}, which is so small that in an Earth-sized block of superconducting aluminum, one would expect to find only two thermally excited quasiparticles. Unfortunately, as we will describe later, observed values of x_{QP} are much larger than expected.

So what happens if quasiparticles are present in a superconducting circuit? In bulk superconductors, they're responsible for finite AC dissipation proportional to x_{QP}. In qubit circuits comprising Josephson junctions, the situation is more complex. When a quasiparticle tunnels from one side of a junction to the other, its coupling to the phase difference across that junction makes it possible for the quasiparticle to absorb energy from the qubit, causing the qubit to decay. Similar to the dissipative response of bulk superconductors, the decay rate is proportional to x_{QP}. Even if the quasiparticle does not absorb energy, when it tunnels it can make the qubit frequency fluctuate, which leads to dephasing and a reduction of the qubit's coherence time. Both energy decay and dephasing originate from the dependence of the tunneling amplitude on the phase difference and have been investigated in a number of theoretical and experimental works (see references 3 and 4 and references therein).

The decoherence mechanisms are generic to any super-

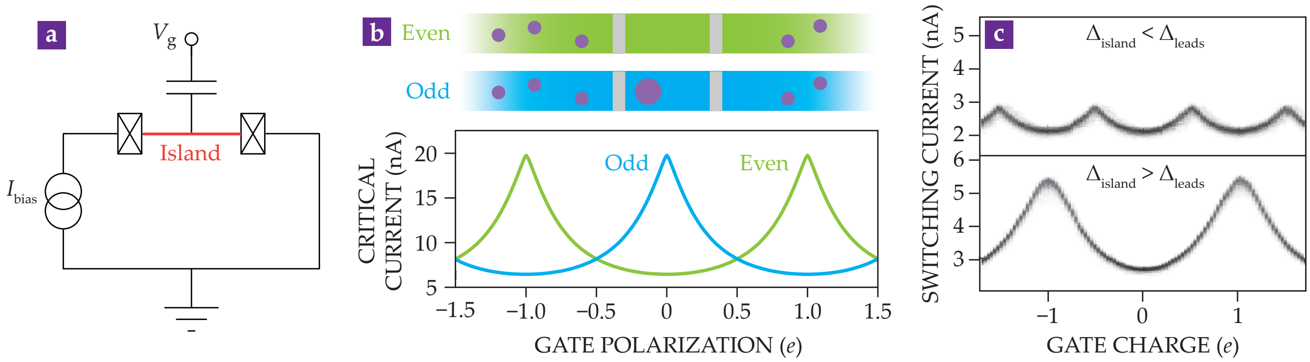


FIGURE 2. SUPERCONDUCTING CIRCUIT. (a) A Cooper-pair transistor circuit features two small Josephson junctions that isolate a submicron-scale superconducting island (red). **(b)** An odd parity state (blue) corresponds to an excess electron on the island, and an even parity state (green), to no excess electron. The effective critical current through the island modulates with an applied gate voltage V_g that corresponds to a change in the energy cost of placing additional Cooper pairs on the island. **(c)** The switching current, which is closely related to the critical current, has a value at a given gate voltage that reflects the presence or absence of quasiparticles poisoning the island charge state. The dips at ±1*e* indicate that single quasiparticles occupy the island more often than not (top). The opposite (bottom) is true when the relative gap energy Δ of the superconducting island and superconducting lead is inverted. (Adapted from ref. 5.)

conducting qubit made with junctions, but different qubit designs have different sensitivities. In fact, qubits with junctions embedded in a superconducting loop can be tuned by threading a magnetic flux through that loop, and the sensitivity to quasiparticles can be suppressed at particular flux values known as sweet spots. The suppression is an interference effect that manifests the nature of quasiparticles as a coherent superposition of electron- and hole-like excitations. At the sweet spots, the sensitivity to flux noise is also minimized, making them by far the preferred operating point for such qubits.

The quasiparticle mystery

As mentioned above, no thermally excited quasiparticles should be present at temperatures sufficiently below T_c. Aluminum circuits with T_c = 1.2 K and at dilution refrigerator temperatures of 10 mK should be completely free of quasiparticles. So why worry about them at all?

In the 1990s several groups studied a class of superconducting charge-sensitive circuits that leveraged the so-called Coulomb blockade effect. In those devices, one or more submicrometer-scale superconducting islands were weakly coupled to connected electrodes by Josephson junctions. Importantly, the small size of the islands and junctions—typically no larger than 100 nm × 100 nm—fixed the islands' total capacitance C_Σ to less than a femtofarad. At that level, the corresponding charging energy for adding a single Cooper pair, $E_C = 2e^2/C_\Sigma$, where e is the electron charge, could easily exceed 10^{-23} J, or 1 K in temperature units.

In that parameter regime, the critical current and other electronic properties were sensitive to the addition or subtraction of single Cooper pairs and quasiparticles. Although quasiparticles do not have definite charge, when they tunnel on or off a superconducting island, the total charge on that island is shifted by the discrete value ±*e*.

One of the simplest Coulomb blockade circuits is the single Cooper-pair transistor.[5] As shown in figure 2, the device has two small Josephson junctions that isolate a single superconducting island from superconducting leads, and a capacitively coupled gate electrode is placed nearby. In that configuration, the two junctions behave effectively as a single Josephson junction. Its critical current—the maximum current that the junction can carry while keeping the voltage across the junction close to zero—modulates with an applied gate voltage. Ideally, the modulation is a 2*e*-periodic function of the gate charge $q_g = C_g V_g$ (where C_g is the gate capacitance to the island, and V_g is the gate voltage) and reflects the size of the Cooper-pair charge itself. As noted above, the presence of quasiparticles in the leads provides a source for single electrons to tunnel onto the island and offset the island's charge by an electron, which concomitantly shifts the current modulation by 1*e*.

Many experimentalists therefore regarded a 1*e*-periodic modulation to be indicative of the presence of quasiparticles. Indeed, one could turn a 2*e*-periodic modulation into a 1*e*-periodic modulation just by heating up the device to a few hundred millikelvin to create an abundance of thermally generated quasiparticles. It was common, however, to see 1*e*-periodic modulation at much lower temperatures, even when controlling for other known causes of the behavior. It's known as quasiparticle poisoning, and its sporadic presence in some, but not all, devices was one of the first indications that the physics of quasiparticles was not fully understood.

Using a higher-speed DC measurement technique in the early 2000s, one of us (Aumentado) found evidence for quasiparticles at dilution-refrigerator temperatures, even in 2*e*-periodic devices. The results showed that the tunneling of nonequilibrium quasiparticles on and off the island was sensitive to both gate voltage and the relative gap energies of the island and leads. Single Cooper-pair transistors share many things in common with today's superconducting qubit circuits, including the junction sizes and material choice of aluminum, and perhaps that's why it's not surprising that the basic phenomenon of nonequilibrium quasiparticle poisoning has persisted to the present day.

To probe the dynamics of nonequilibrium quasiparticles in superconducting qubits and test our understanding of quasiparticle poisoning, researchers have used many approaches over the years. For example, one can purposely add quasiparticles by increasing the system's temperature and then measuring such properties as the relaxation time T_1 (typically tens to hundreds of microseconds) and the qubit frequency ω_{10} (a few

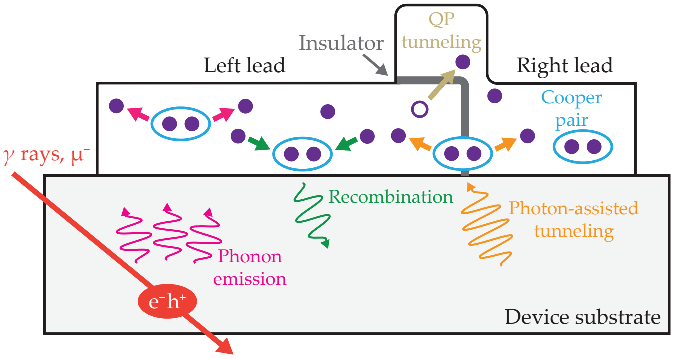


FIGURE 3. QUASIPARTICLE DYNAMICS. A Josephson junction, formed by a superconductor-insulator-superconductor heterostructure, is shown in cross section. Quasiparticles (purple) can undergo various inelastic processes. Some tunnel across the Josephson junction (yellow) and others are generated during photon-assisted tunneling of Cooper pairs (orange). Both processes can cause energy exchange between the quasiparticles and a qubit formed in part from the junction. Ionizing radiation can create in the substrate electron–hole pairs (red), which emit showers of phonons (pink) as they relax. Phonons with an energy of 2Δ or greater are sufficiently energetic to break Cooper pairs; freshly created quasiparticles in the device then lead to spatiotemporally correlated errors. Quasiparticles can also recombine and emit a phonon with energy greater than 2Δ (dark green).

gigahertz). Both those properties decrease when quasiparticles are present.[6]

Alternatively, nonequilibrium quasiparticles can be injected directly without raising the system temperature, and the expected relation between changes in T_1 and ω_{10} can be checked.[7] In fact, researchers have exploited the proportionality between $1/T_1$ and the quasiparticle density x_{QP} to monitor the dynamics of x_{QP}, and they have assessed to what extent quasiparticles were trapped by supercurrent vortices.[8] Such experiments also make it possible to place bounds on the density of nonequilibrium quasiparticles and to estimate their generation rate.

A more direct measure of quasiparticle effects in qubits is similar to the initial observations of 1e periodicity in single Cooper-pair transistors.[5] By explicitly reintroducing some charge sensitivity into a transmon circuit, researchers detected quasiparticle-induced errors via a correlated change in the odd–even "charge parity" of the circuit over a time τ_{QP} (see reference 9 and references therein). From those experiments, it's clear that modern-day superconducting qubits are still plagued by nonequilibrium quasiparticle poisoning.

Sources of nonequilibrium quasiparticles

Once physicists accepted that nonequilibrium quasiparticles were present in their superconducting devices, a simple question remained: Why? The answer boils down to the erroneous assumption that everything a qubit "sees" is perfectly isolated from the outside world and well-thermalized to the coldest stage of the cryostat. For low-noise experiments with superconducting qubits, researchers take a lot of care to filter and shield any unwanted noise. But qubits aren't ever completely sheltered. All it takes to produce a pair of quasiparticles in an otherwise isolated superconductor is an excitation with an energy greater than 2Δ, which for commonly used thin aluminum films corresponds to approximately 100 GHz, 5 K, or 400 μeV, depending on the preferred choice of units. That's not a lot of energy!

The many years that researchers have spent developing superconducting detectors have led to valuable insights into the dynamics of nonequilibrium quasiparticles. Figure 3 summarizes how all sorts of bad actors, including stray IR photons, mechanical vibrations of the device, and—most troubling of all—ionizing radiation from radioactive decay products and cosmic-ray secondary particles generate quasiparticles in qubits.

IR photons can leak into the experimental region of a cryostat, despite the best attempts to block or shield from them. Many popular cryogenic systems, including dilution refrigerators, consist of multiple temperature stages. Similar to a set of nested matryoshka dolls, a metal shield at each stage protects the next from the surrounding, hotter stage (see figure 4). The innermost shield should be thermalized to the lowest-temperature stage of the cryostat. Experiments with superconducting resonators, however, indicated that more shielding was needed: Some photons from higher-temperature stages can get through and reduce device performance.[10] Coating the experiment with IR-absorbing material is one remedy. It's the same principle that's used when painting stealth aircraft.

Researchers recently discovered that the qubit itself can act as an antenna that enhances the production of quasiparticles via absorption of IR radiation.[11] The absorption process is localized at the Josephson junctions of a qubit circuit; in addition to qubit relaxation, the process can explain recent observations of especially large qubit excitation rates.[9] Experiments have since demonstrated that the process does indeed contribute to quasiparticle generation and qubit excitation and that the process can be suppressed by improved filtering of the microwave lines feeding signals to the qubits and by proper design of the qubit and its surroundings.[9,12] Those improvements can lengthen by several orders of magnitude the time between quasiparticle tunneling events, from shorter than a millisecond to longer than a second.

Ionizing radiation is known to also produce quasiparticles in superconducting devices, and in many cases that's the desired effect. So-called pair-breaking detectors, such as microwave kinetic inductance detectors and transition-edge sensors, operate on the principle that ionizing radiation and other excitations deposit large amounts of energy into the crystalline device substrate in the form of ionized charge carriers and showers of high-energy phonons. In superconducting detectors, the phonons can produce quasiparticles, whose presence is inferred from a change in an observable parameter, such as kinetic inductance or critical current.

Although superconducting qubits are similar in construction to those types of detectors, it was only in hindsight that researchers realized that superconducting qubits could also act as detectors of ionizing radiation, with detection events translating into computational errors. Ionizing radiation reduces the performance of qubits.[13] Some of it, primarily γ rays, can be shielded by lead, but to cut down on the flux of pesky cosmic-ray muons, one needs to use the overburden of Earth's crust or to go deep underwater.[14]

The mechanism of quasiparticle production via cosmic-ray muons is particularly worrisome because about every 10 seconds a muon can generate bursts of quasiparticles throughout a device and knock out many nearby qubits simultaneously.[15]

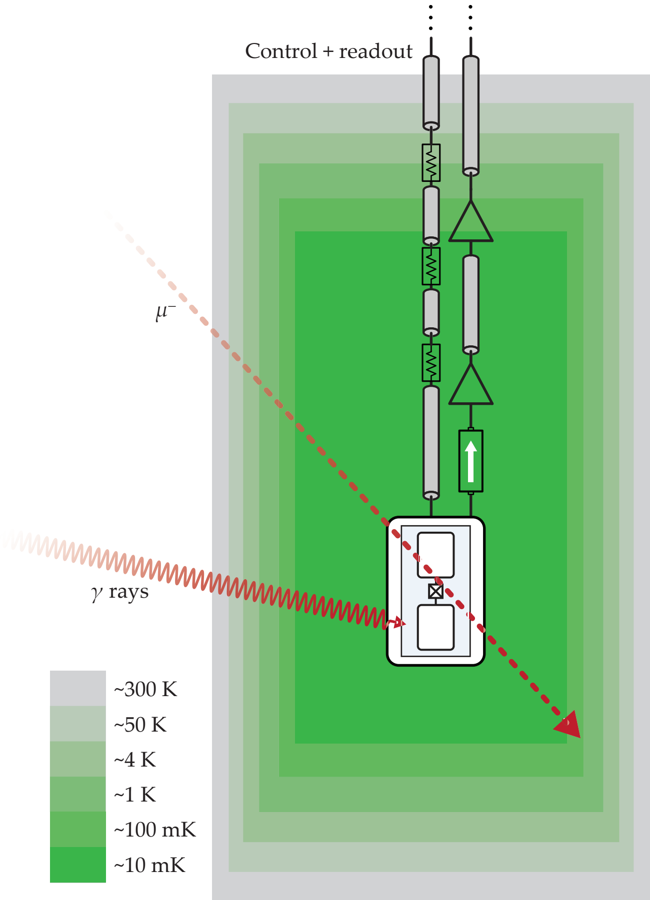


FIGURE 4. SUPERCONDUCTING QUBIT EXPERIMENTS often use dilution refrigerators with nested temperature stages. Each stage includes a metallic shield that blocks blackbody radiation from higher-temperature stages. Gamma rays and cosmic-ray muons, however, can penetrate through that shielding, sometimes hitting the superconducting quantum processor and creating spatiotemporally correlated, quasiparticle-induced errors.

Similar bursts were recently linked to mechanical relaxation of superconducting devices over the time scale of days. The link could explain an earlier observation of a slow decay in the generation rate over the course of an experiment. Those types of quasiparticle-induced spatiotemporally correlated errors are difficult to deal with in many quantum error-correction schemes, although they can be addressed if they're detected independently and if qubits likely to have been affected by errors can be excluded from further computation.[16]

Toward robust quantum computing

Qubit performance has improved by several orders of magnitude in the 25 years since the first demonstration of coherence in a superconducting qubit, but there is still a long road ahead. The consensus in the research community is that quantum error-correction techniques will be necessary to maintain complex multiqubit-state information for the duration of a useful computation. In such schemes, logical qubits are encoded in the combined state of many error-prone qubits, and higher error rates translate into stricter requirements on the total number of physical qubits.

An underlying assumption typical of quantum error-correction schemes is that physical errors are random. Using that thinking, researchers have steadily chipped away at the background population of nonequilibrium quasiparticles and suppressed their steady-state contribution to qubit errors to a sufficient level over time. But that assumption is violated by the aforementioned error bursts that arise from quasiparticles generated by ionizing radiation.

Luckily, there are many proposed—and some demonstrated—paths toward mitigating catastrophic error bursts. Having quasiparticles around is ok, so long as they don't tunnel across a qubit's Josephson junction. That could be achieved by using a superconductor for the ground plane with a smaller energy gap than the qubit superconductor or by adding normal-metal islands to the back of the chip.[17] Those design changes bring the energy of the phonons generated by radiation hits to below the gap of the qubit material, so that they cannot break Cooper pairs anymore. The few quasiparticles that are still generated in the qubit bulk can be kept away from the qubit's junctions by employing quasiparticle traps[9,18] or blocked from tunneling at the junctions via gap engineering.[5]

While those "on-chip" techniques are effective for many sources of quasiparticles, pesky cosmic-ray secondary particles such as muons are not effectively shielded except by massive amounts of material, which has led some scientists to suggest that underground facilities are critical to avoiding spatiotemporally correlated error bursts. Luckily for experimentalists who enjoy sunlight, there is hope that on-chip mitigation strategies could be combined with tungsten or lead shielding to provide sufficient protection. But such radiation-hardened superconducting qubits have yet to be fully demonstrated.

Nonequilibrium quasiparticles might sound like a bogeyman lurking in the shadows of superconducting quantum computing efforts, but they are just another item in the list of engineering and scientific challenges that must be met to make quantum computing a robust reality. There are many reasons to be optimistic: Recent research efforts have given more insight into quasiparticles' generation mechanisms and have provided a clear direction for future mitigation efforts.

REFERENCES

1. P. Krantz et al., *Appl. Phys. Rev.* **6**, 021318 (2019).
2. I. Siddiqi, *Nat. Rev. Mater.* **6**, 875 (2021).
3. L. I. Glazman, G. Catelani, *SciPost Phys. Lect. Notes* (2021), doi:10.21468/SciPostPhysLectNotes.31.
4. G. Catelani, J. P. Pekola, *Mater. Quantum Technol.* **2**, 013001 (2022).
5. J. Aumentado et al., *Phys. Rev. Lett.* **92**, 066802 (2004).
6. G. Catelani et al., *Phys. Rev. Lett.* **106**, 077002 (2011).
7. M. Lenander et al., *Phys. Rev. B* **84**, 024501 (2011).
8. C. Wang et al., *Nat. Commun.* **5**, 5836 (2014).
9. K. Serniak, "Nonequilibrium quasiparticles in superconducting qubits," PhD thesis, Yale U. (2019).
10. R. T. Gordon et al., *Appl. Phys. Lett.* **120**, 074002 (2022).
11. M. Houzet et al., *Phys. Rev. Lett.* **123**, 107704 (2019).
12. T. Connolly et al., https://arxiv.org/abs/2302.12330.
13. A. P. Vepsäläinen et al., *Nature* **584**, 551 (2020).
14. L. Cardani et al., *Nat. Commun.* **12**, 2733 (2021).
15. M. McEwen et al., *Nat. Phys.* **18**, 107 (2022).
16. Q. Xu et al., *Phys. Rev. Lett.* **129**, 240502 (2022).
17. V. Iaia et al., *Nat. Commun.* **13**, 6425 (2022).
18. R.-P. Riwar et al., *Phys. Rev. B* **94**, 104516 (2016).

PT

The Daniel K. Inouye Solar Telescope, on the Haleakalā volcano on the Hawaiian island of Maui. (Courtesy of NSO/AURA/NSF, CC BY 4.0.)

Holly Gilbert is the director of the High Altitude Observatory at the National Center for Atmospheric Research in Boulder, Colorado. She previously served as the director of the heliophysics science division at NASA's Goddard Space Flight Center.

ADVANCES IN SOLAR TELESCOPES

Holly Gilbert

Even as our understanding of the Sun has grown, many fundamental questions remain—some of which have big implications for life on Earth.

Our Sun has been an object of mystery and study for as long as humans have been around to squint at it in wonder. Sunspot observations have been recorded for over 400 years, and eclipses were first documented thousands of years ago. Although the Sun is no longer a complete mystery, the way we observe it continues to develop, as continued technological advances uncover new mysteries for future generations to solve.

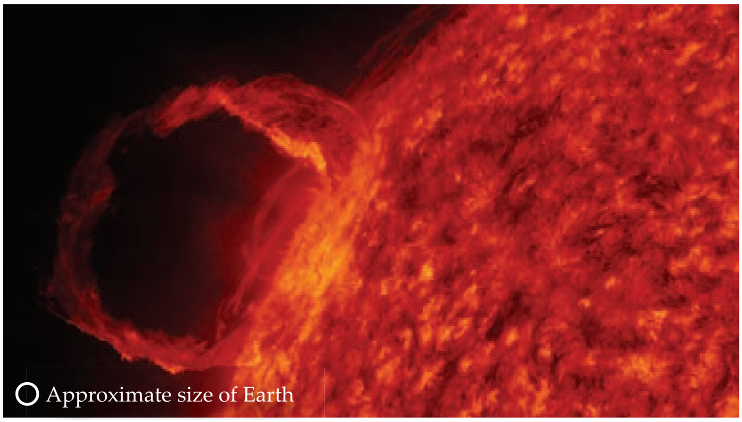


FIGURE 1. LARGE ARCS OF SOLAR MATERIAL, or prominences, are common features of an active sun. They were only discovered once astronomers viewed the Sun through telescopes. The *Solar Dynamics Observatory* now provides high-resolution monitoring of such solar activity. (Courtesy of NASA's Goddard Space Flight Center Scientific Visualization Studio and the *Solar Dynamics Observatory*.)

The invention of telescopes facilitated more detailed study of sunspots and the identification of exotic structures suspended above the solar surface (see figure 1). More recently, the space age brought about a new era of solar observations, leading to groundbreaking discoveries about solar magnetism and its reach. Launched in the 1970s, *Skylab* was the first space station, and although not designated to be a solar mission, it opened our eyes to large eruptions originating in the outer layer of the solar atmosphere, called the corona. Since the launch of the *Solar and Heliospheric Observatory* (*SOHO*) in the 1990s, a joint project between the European Space Agency and NASA, the trajectory for solar physics discoveries has swiftly escalated. Despite advances, many fundamental questions remain regarding all regions of the Sun, including the solar interior and the storms released in the outer solar atmosphere.

Pesky problems in solar physics

The Sun experiences an 11-year activity cycle. During its peak, called the solar maximum, the number of sunspots and related eruptions increases significantly, at times averaging three storms a day. In the solar-minimum phase, there can be no sunspots for long periods of time. The Maunder Minimum was a seven-decade period in the 17th and 18th centuries during which sunspots became extremely rare. From 1672 to 1699, observations revealed fewer than 50 sunspots, compared with the thousands of sunspots typically seen over similar time spans. Astronomers still don't know what causes the 11-year solar cycle, let alone what led to the unusual lack of sunspots during the Maunder Minimum.

George Ellery Hale (1868–1938) and collaborators determined that sunspots and their 11-year cycle were of a magnetic nature in the beginning of the 20th century. Although we now know plasma flows and magnetism are at the core of that phenomena, the solar dynamo driving those fluctuations is still being studied and modeled. A missing piece of information that would help advance the modeling efforts is the Sun's global magnetic field, especially data covering the polar regions. To date, the polar regions of the Sun have never been observed since the poles are not easily seen from Earth.

Solar flares are also not fully understood. Among the most explosive phenomena in the solar system, they were discovered in 1859 during the landmark Carrington event.[1] During his regular monitoring of sunspots, amateur astronomer Richard Carrington observed rapidly brightening patches of light near a sunspot group; the changes in brightness were visible without any advanced technology. That flare was associated with strong effects seen on Earth shortly after: Telegraphs caught fire, and auroras were seen as far south as the Caribbean. It proved that solar events could impact Earth. Being able to predict solar flares of that scale is more important now than ever before because of our increased reliance on technology.

Coronal mass ejections (CMEs) are another solar phenomenon powered by magnetic fields. Associated with flares and driven by magnetic reconfigurations, they were first observed from the ground in 1971, and their dynamics were elucidated by *Skylab* observations a few years later. Magnetic field interactions, especially in the corona, are at the root of eruptions and flares, but exactly how and where fields interact is still being debated. There are several models that connect magnetic topologies to observations of erupting materials and fields. The coronal magnetic field is also the energy source and organizing principle behind solar-wind acceleration[2] and coronal heating, two related and complex physical phenomena.

The coronal-heating problem is perhaps the longest standing, most frustrating issue yet to be resolved in the solar-physics community.[3] In daily life, you expect that you will feel cooler as you move farther away from a heat source. But the Sun's corona—despite being the outermost layer of the solar atmosphere—is a million kelvin hotter than the Sun's visible surface, the photosphere (see figure 2). Moreover, that temperature gradient occurs across a very short distance of about 100 km, called the transition region. Recent satellite missions (see the article by Leonardo Di G. Sigalotti and Fidel Cruz, Physics Today, April 2023, page 34) have partly solved the mystery of why the outer layer is the hottest. Once energy from the core reaches the surface, it is transported through the chromosphere, transition region, and corona via magnetohydrodynamic waves or magnetic-field-line braiding and subsequent reconnection.

Magnetic reconnection occurs when oppositely directed magnetic field lines in a plasma break and rearrange. That process releases a large amount of energy that can heat the plasma and accelerate energetic particles. Magnetohydrodynamic waves, the other method of energy transport, access energy in the subsurface convection zone and transport it through plasmas. In the corona, to heat the plasma that wave energy must dissipate a nontrivial amount of heat because of the behavior of damping time scales in plasmas. It is widely accepted that dynamic magnetic fields play a fundamental role in those processes, but astronomers don't yet have enough information to accurately model the complex nature and relative role they play in heating the solar atmosphere.

Obtaining measurements of the coronal magnetic field vector and inferring properties of the fields are essential steps in understanding the physics of such phenomena. Solving those outstanding problems—or at least significantly advancing our understanding of them—is of particular importance in the con-

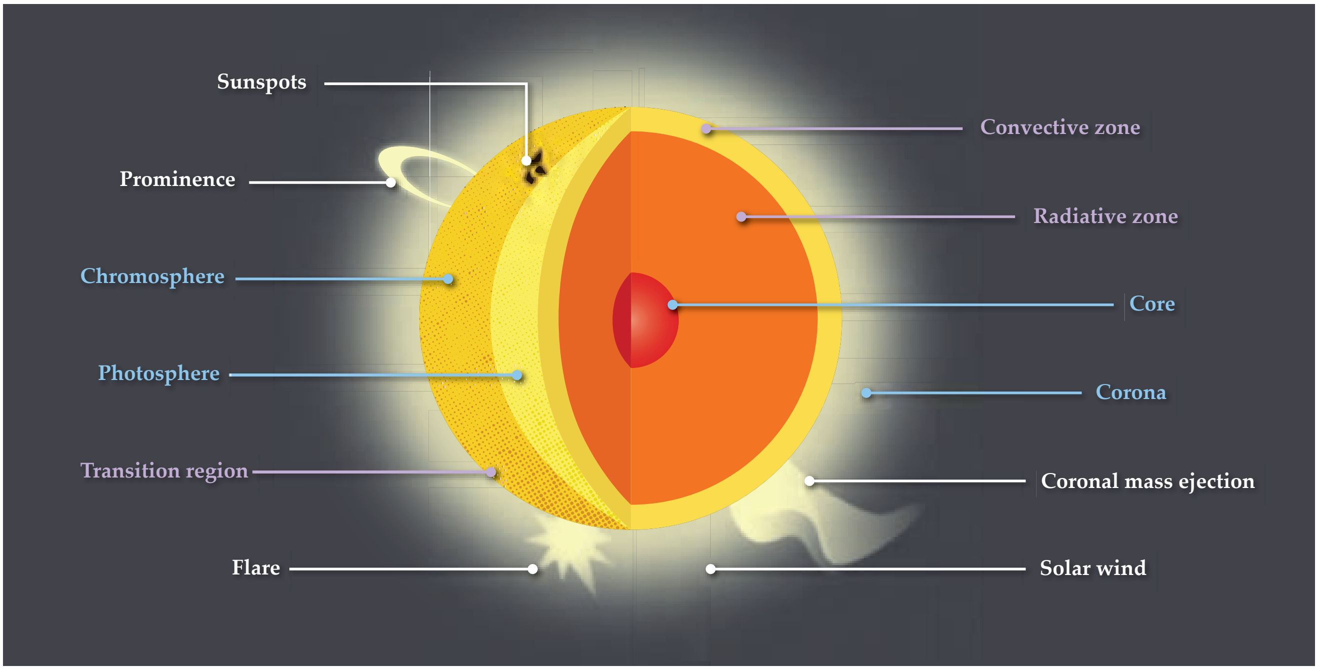


FIGURE 2. THE OUTER LAYERS OF THE SUN are the easiest to study, but even those still hide mysteries. The causes of surface phenomena (in white) such as flares and coronal mass ejections lie in the interior layers, which are harder to access. (European Space Agency image adapted by Abigail Malate.)

text of space weather. Our strong technology dependence in the modern era makes us increasingly vulnerable to the effects of Sun–Earth interactions during solar storms, especially energetic charged particles and magnetic fields carried great distances in CMEs. Those interactions affect the power grid, communications and other satellites, airlines flying over the poles, navigation, and astronaut health, making space weather a subject of national priority. As recently as 2020, PROSWIFT—a US bill to promote research to improve space weather forecasting—was signed into law.

Observational needs

A basic tenet of observational science is the constant need for improved spatial (and in the case of solar physics, temporal) resolution. You'd be hard-pressed to find a scientist who wouldn't advocate for it, especially in the solar-physics community. The physics and dynamics of the Sun necessitate a multiscale problem that can only successfully be understood with improvements to observations of solar activity. With each new mission or telescope, improved spatial and temporal resolution leads to new discoveries. Often the field of view (FOV) must be compromised to accommodate the extra resolution capability. Therefore, of equal importance is the existence of comprehensive, full-Sun images to provide the large contextual view. Large FOV data sets capture the global consequences of processes that occur on short time scales and small spatial scales.

Because the Sun is a three-dimensional sphere, multiple viewpoints are essential for obtaining accurate information about its inherent magnetic and thermal nature. Helioseismology (see the article by John Harvey, Physics Today, October 1995, page 32) allows a unique view into the solar interior using methods similar to seismology on Earth. A few space missions, such as the *Solar Terrestrial Relations Observatory*, have flown in special orbits designed to collect data from 360° around the Sun. The *Parker Solar Probe* (see the article by Nour E. Raouafi, Physics Today, November 2022, page 28) and *Solar Orbiter* fly extremely close to the Sun—within the solar corona. Other mission concepts are being developed to cover still more regions of the Sun. Together, those types of extended viewpoints provide a global context that is critical to understanding the basic physical processes driving solar activity.

A common theme behind the unsolved solar-physics problems is the need for magnetic field measurements. But those measurements remain difficult to make, and the inferences drawn from existing observations are most mature for the photosphere. In the higher layers of the atmosphere, however, advances continue to be made in inferring the magnetic field through spectropolarimetry, the measurement of polarization as a function of frequency.

Detection methods

Studying light from the Sun involves more than just counting how many photons are detected and in which filters. Coronagraphs, which occult the bright disk of the Sun, allow observations of the diffuse corona. Polarimeters, which measure the polarization of light, can be used to infer information about the magnetic field and composition of the Sun. Both instruments select components of the total light from the Sun to gain a better understanding of its dynamics.

White-light coronagraphs allow glimpses of the coronal plasma and open observational windows to some of the most important causes of space weather: CMEs. Invented by astronomer Bernard Lyot in the 1930s, the coronagraph was innovative due to its aperture—called the Lyot stop—which blocks light diffracted around the entrance aperture. Coronagraphs work by allowing light to enter the telescope aperture as an evenly illuminated source. They require the rejection or suppression of stray light to a very high degree using a properly designed occulter system made up of baffles, stops, an internal occulter, optical coatings, out-of-band light rejection, and a highly polished objective lens.

Figure 3 shows a coronagraph design in which the secondary mirror at the center blocks the light with a lens that subsequently images it. Where a camera or detector would usually record the image, an occulting spot, also called a focal plane mask, is placed instead. That absorbs most of the light from the center of the FOV, while the telescope pupil is reimaged by another lens. The remaining light from the central source is concentrated around the edges of the pupil, forming rings around the edge of the aperture image and the secondary mirror image. The goal is to block as much unwanted light as possible by the time it reaches the detector. The next-generation coronagraphs have the capability to use spectropolarimetry to study flows, waves, density, and magnetic fields.

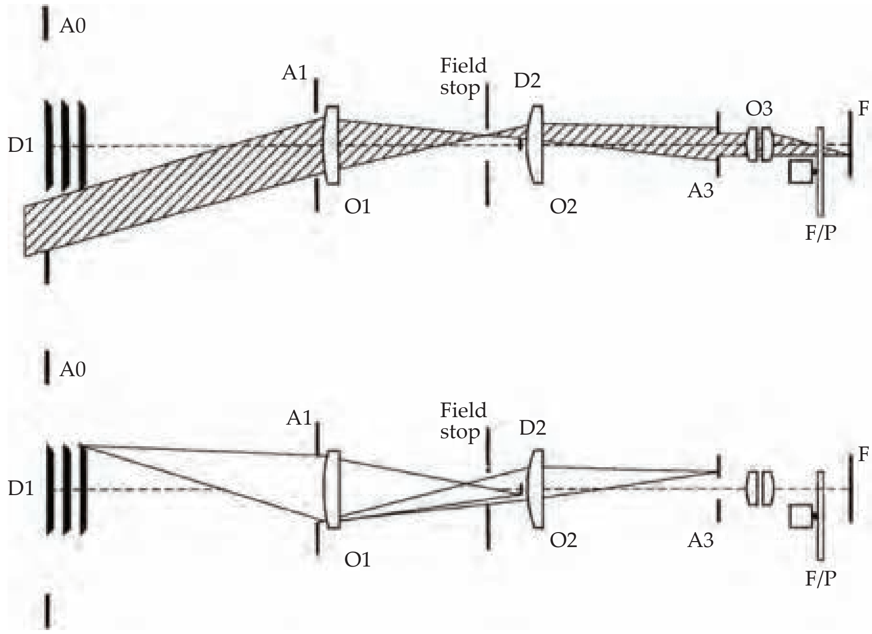


FIGURE 3. THE LARGE ANGLE AND SPECTROMETRIC CORONAGRAPH C3 on the *Solar and Heliospheric Observatory* (*SOHO*) satellite. The top diagram illustrates image formation; the bottom diagram shows how stray light is blocked from reaching the image plane. Sunlight enters from the left. The components shown are the front aperture (A0), external occulter (D1), entrance aperture (A1), objective lens (O1), field stop, internal occulter (D2), field lens (O2), Lyot stop (A3), relay lens with Lyot stop (O3), filter and polarizer wheels and shutter (F/P), and focal plane (F). (Adapted from ref. 10.)

Polarization of light is expected whenever the mechanisms producing the observed radiation—whether radiative or collisional—act in the presence of symmetry-breaking conditions. For example, the scattering of Planckian blackbody radiation in a homogenous and isotropic gas will be unpolarized. The same process, however, will produce polarized light if the gas is illuminated nonisotropically, where the bulk of radiation comes from a preferential direction. Polarized light is seen, among other places, in the scattering of solar-disk light by coronal structures observed off the solar edge.

Another typical symmetry-breaking situation producing polarized light is the radiative emission by atoms in the presence of an ambient magnetic field. In those cases, our ability to model the production of polarized light and interpret its signature in our data allows us to diagnose the thermodynamic and magnetic conditions of the emitting gas. Solar scientists largely rely on theoretical, numerical, and experimental polarimetric tools to diagnose the magnetism of the solar atmosphere.

The characterization of polarized light requires the specification of additional radiation quantities beyond intensity. Two parameters are needed to fully specify the degree and direction of linear polarization. A third parameter specifies the state of circular polarization about the propagation direction. Light detectors, however, are typically only sensitive to intensity signals. Hence, a polarimeter generally processes the incoming polarized light through some variable, birefringent optic system (the modulator) that encodes the polarization information into a finite series of varying intensity signals that can be captured with our detectors after being filtered by a linear polarizer (the analyzer).

Observing from the ground

Despite the availability of satellites, ground-based solar observatories continue to be an important source of solar imagery, quantitative spectroscopy, and polarimetry. Remaining on the ground offers multiple advantages, one of the most obvious being cost efficiency. Going to space is expensive and limits the ability to service and maintain instruments, whereas on the ground, it is easy to physically access the instruments for repairs and upgrades. Moreover, there are limits to the size of the telescope aperture and the complexity of the instruments that can be afforded in space. The largest solar telescope in space is the 50 cm *Hinode*, which is an order of magnitude smaller than the newly functioning ground-based Daniel K. Inouye Solar Telescope (DKIST), discussed below.[4] DKIST offers unprecedented views of the Sun that would be impossible to obtain in space because of cost limitations.

Ground-based white-light coronal instruments have been operating at the High Altitude Observatory site in Climax, Colorado, from 1956 to 1963 and at NSF's Mauna Loa Solar Observatory (MLSO), located on the island of Hawaii, since 1965. Although limited by weather and the nighttime hours, a series of coronameters and coronagraphs (each instrument an improvement over the previous one) has provided important information on the innermost part of the corona and on the origins of CMEs. According to models, those large eruptions are often initiated by magnetic reconnection in the low corona, so observing close to the solar surface is advantageous, something that the design of the MLSO white-light coronagraph enables (see figure 4). The Coronal Multi-channel Polarimeter (CoMP) located at MLSO and its upgraded follow-on, UCoMP, are another type of ground-based coronagraph. UCoMP offers a unique combination of magnetic diagnostics, thermal information, and Doppler motions because of its sampling of multiple coronal lines, and its new filter is designed for a very broad bandpass that covers a range of 5000 Å.

UCoMP is a prototype for a larger telescope being developed as part of the Coronal Solar Magnetism Observatory, a proposed suite of complementary ground-based instruments (funding and location to be determined) designed to study magnetic fields and plasma conditions in the Sun's atmosphere.[5] The primary instrument is a large coronagraph (the next-generation UCoMP) that offers better resolution, polarization measurements, and the ability to obtain line-of-sight magnetic fields (UCoMP only provides plane-of-sky magnetic fields). The suite's supporting instrumentation—a chromosphere magnetometer and a white-light coronagraph—measure magnetic fields in the Sun's chromosphere and the density of electrons in the corona. Together, those provide new tools such as vector-field measurements and invaluable predictive clues about damaging solar events.

The larger a telescope is, the more photons it can collect and the higher resolution it can achieve. NSF's DKIST on the Haleakalā volcano on Maui (see page 40) is now the largest solar telescope in the world. It was designed as a coronagraph that enables coronal spectroscopy and magnetometry. It collects more sunlight than any solar telescope and provides the best resolved and sharpest images of the Sun. DKIST's innovative technology includes an off-axis optical system with a large (4 m) primary mirror, active and adaptive optics, advanced optical and IR instruments, and versatile light-distribution optics that facilitate the simultaneous use of multiple instruments.[4] DKIST's instrument capabilities are centered on spectropolarimetry. The ability to precisely measure the magnetic field throughout the solar atmosphere, including the corona, enables DKIST to address basic research aspects of space weather. DKIST will operate for two solar cycles and conduct community-proposal-driven observations.

An even more powerful tool being planned is a network of ground-based instruments at various longitudes designed to obtain wide coverage of observations. NSF's Global Oscillation Network Group (GONG) has served a vital role in comprehensive measurements of solar oscillations and magnetic field measurements in the photosphere since it was built in the mid 1990s. The six-site network (Australia, India, Canary Islands, Chile, California, Hawaii) was not designed with space weather in mind; that requires better sensitivity and new capabilities. The next-generation GONG is being developed to fill that critical gap once GONG's lifetime expires sometime in the 2030s, owing to aging instrumentation and the difficulty of replacing old hardware that no longer exists for purchase. Through a partnership of NSF institutions (the National Solar Observatory and National Center for Atmospheric Research) and potentially other government agencies, the next-generation GONG will install at each site some combination of spectropolarimeters for the precise measurements of solar magnetic fields at multiple heights, coronagraphs capable of monitoring the violent ejecta of magnetized plasma from the Sun's atmosphere and determining coronal magnetic topologies and plasma properties, and instruments for Doppler-velocity measurements required for helioseismology studies.

Another potential ground-based network of solar instrumentation is being explored by a group of international observatories. They are currently carrying out a preliminary design study of a synoptic solar-observing facility called the Solar Physics Research Integrated Network Group, funded by the European Union.[6] The data products will include images with arcsecond resolution in multiple wavelengths, synoptic data including vector magnetic fields and surface velocity, and observations of flares and other transient events.

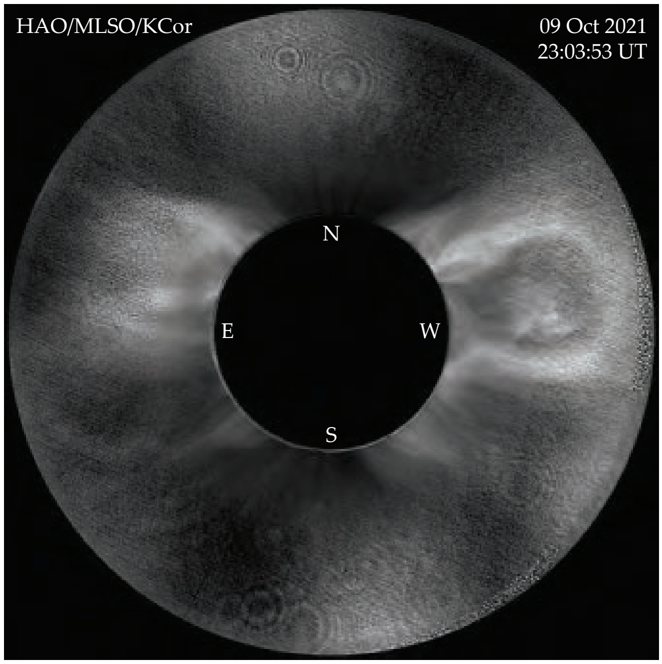


FIGURE 4. A WHITE-LIGHT IMAGE from Mauna Loa Solar Observatory's K-coronagraph shows a coronal mass ejection on the right side of the masked solar disk. The eruption occurred on 9 October 2021. A video of the coronal mass ejection is available at the online version of this article. (Courtesy of HAO/MLSO.)

Other methods based on radio emissions provide constraints on the magnetic field in the corona by utilizing multiwavelength observations that are sensitive to both thermal plasma and nonthermal particles in addition to solar magnetic fields. The science of solar radio astronomy is many decades old. The Owens Valley Solar Array, established in 1979 and located near Big Pine, California, performs radio interferometry at many frequencies and conducts microwave imaging spectroscopy. It can detect thermal radiation in the chromosphere and corona and nonthermal radiation from solar flares. The Frequency Agile Solar Radiotelescope is a next-generation radio telescope that is being proposed to NSF (location and funding to be determined) and is specifically for solar observations.[7,8] It's designed to perform Fourier synthesis imaging, exploiting the Fourier-transform relationship between the quantity measured by an interferometer and the radio-brightness distribution, and will combine ultrawide frequency coverage, high spectral and time resolution, and excellent image quality. That approach to solar observing offers a unique capability of measuring coronal fields against the solar disk.

Observing from space

The next generation of ground-based telescopes provides spectacular opportunities for new solar data and a chance at making exciting discoveries, but in the present age, ground-based data are made better when combined with space-based data. The obvious downside to observing from the ground is the

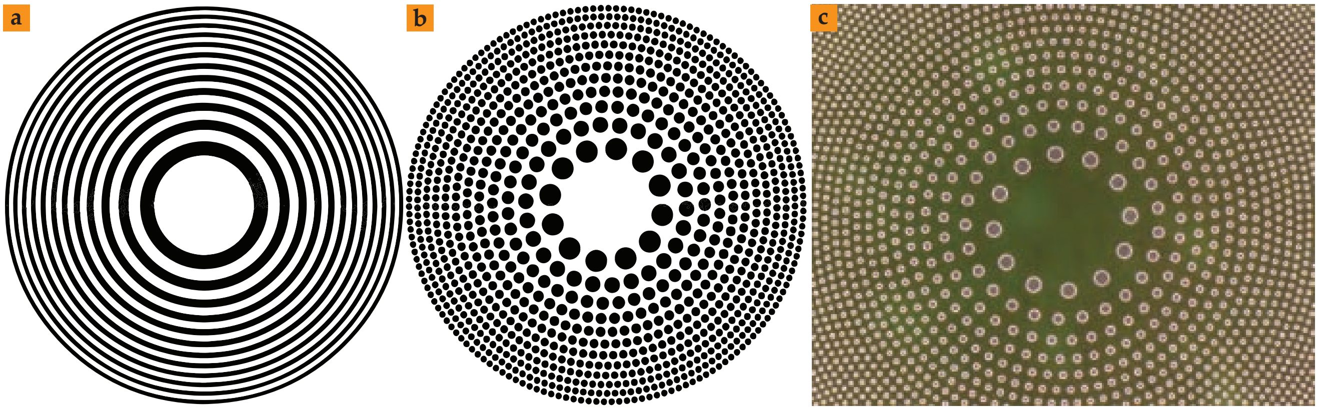


FIGURE 5. ZONE PLATES are one way to image the Sun near the diffraction limit. **(a)** A classical Fresnel zone plate has concentric rings with decreasing widths. **(b)** The photon-sieve variant has concentric rings of circles with decreasing radii. **(c)** A fabricated photon sieve needs to be precise to achieve the high-resolution images desired. The one seen here has holes as small as 2 μm in diameter. (Panels a and b courtesy of Adrian Daw; panel c courtesy of Kevin Denis.)

inability to observe more than what weather and daytime permit. Also, an important part of the electromagnetic spectrum—including the extreme ultraviolet (EUV) and x-ray regions, both important to seeing the dynamic and explosive activity on the Sun—is not accessible through Earth's atmosphere.

The space age brought with it groundbreaking discoveries in solar physics. One such discovery was *Skylab*'s sequences of images of CME dynamics in the 1970s. Multiwavelength observations, including those only accessible from space, opened windows into the various layers of the solar atmosphere. The limitations of space-based observations include cost and challenging access for equipment repairs. Yet the pros far outweigh the cons, and we have become incredibly adept at building resilient spacecraft and instrumentation, partly due to rigorous processes involved in delivering a proof of concept. Prior to being commissioned for spaceflight, many solar instruments are flown on suborbital sounding rockets to stress test them or on balloons. The long process of getting instrumentation to an appropriate level of technical readiness for space somewhat guarantees low risk and high success rates.

Multiple solar coronagraphs have flown on suborbital platforms since the 1960s, and the advances within the last couple of decades have significantly contributed to better forecasting capabilities for space weather. Because of the stringent size and weight limitations of spaceflight, designers of spaceborne coronagraphs face difficult challenges when trying to eliminate stray light. Those instruments employ multiple external occulters with shapes optimized to minimize the effects of scattered light. For researchers to view the low corona, the external occulters must be located at a very large distance from the objective lens. That makes it challenging for space-based externally occulted coronagraphs to match the FOV of the ground-based K-coronagraph at MLSO, about 1.05 to 3 solar radii.

Space-based solar coronagraphs that use multiple occulters, however, can get a better signal-to-noise ratio farther away from the solar disk and are capable of extending the outer FOV to large distances away from the solar surface; *SOHO*'s Large Angle and Spectrometric Coronagraph C3, for example, extends out to 33 solar radii.[9] Moreover, the almost-continuous observing facilitates the forecasting of Earth-directed CMEs and other activity. The newest conceptual coronagraphs designed for space explore a range of the spectrum not accessible to the ground, offering the potential of obtaining magnetic field measurements through spectropolarimetry. Some newer concepts separate the occulter from the telescope, introducing unique challenges with stability and jitter. Proba-3, a European Space Agency mission planned to launch in 2024, will feature two spacecraft—one an occulter and one a coronagraph—flying in formation 144 m apart, with millimeter accuracy, to create an externally occulted Lyot-style white-light solar coronagraph.

The improved technology, and subsequent rise in popularity, of CubeSats has provided some flexibility in flying instrumentation. CubeSats are small satellites that use a standard size and form factor. The field of heliophysics has been increasingly capitalizing on the more cost-efficient access to space they offer. CubeSats, however, require small instruments, so there has also been a push toward the miniaturization of hardware. Some observing capabilities currently cannot be shrunk beyond a certain threshold, but innovative coronagraphs have decreased in size to be flown with other instruments. Small coronagraphs can utilize folding mirrors or a boom to deploy the occulter and thereby reduce the size of the instrument volume on the spacecraft. That is particularly useful when sending instruments to nontypical orbits in space. Having observations from other viewpoints off the Sun–Earth line of sight is incredibly useful for seeing CMEs directed at Earth and obtaining more accurate information about their speed and location.

Observing only from the Sun–Earth line forces us to look at the Sun as a 2D object, and we miss out on some important activity on its backside and have incomplete information on the magnetic field, which leads to inaccuracies in models of the corona and solar wind. A 360°, synoptic view is ideal. A constellation of small satellites deployed around the Sun and over its poles could get us closer to that ideal. This would also help us understand fundamental and universal processes that apply to other stars.

An unconventional look forward

Considerable indirect evidence suggests that the million-degree solar corona is heated by releases of energy on scales smaller than 100 km, which is about 0.15 arcseconds as seen from Earth. High-temperature solar plasma emits radiation primarily at EUV and x-ray wavelengths, which must be observed from space. But no conventional reflecting telescope has been fabricated with the extreme accuracy and smoothness necessary to achieve diffraction-limited imaging at those wavelengths. EUV wavelengths would require subnanometer accuracy, which is extremely difficult to achieve for meter-scale mirrors. A flat diffractive optic, such as a Fresnel zone plate (see figure 5), has been demonstrated to produce nearly diffraction-limited images with a tolerance at least an order of magnitude looser than conventional optics.

A conventional optical surface such as a mirror or lens is a powerful way to bend light rays. But each surface must have just the right tilt with respect to each incoming ray to produce a focal point. A zone plate can be pictured as transforming each incoming ray not into a single deflected ray but instead into a cone of outgoing rays, only some of which interfere constructively to produce a focal point. An EUV space observatory with a diffractive optic of modest aperture (less than 1 m in diameter) can probe the small spatial scales at which coronal heating is believed to take place. Recent progress suggests a suitable diffractive optic should be available in the near future.

A variant of a zone plate is the photon sieve. A photon sieve 80 mm in diameter has been shown in the lab to produce nearly diffraction-limited EUV images. A mission concept in development calls for 170-mm-diameter sieves to achieve the angular resolution required for its scientific goals. The most challenging aspect of that coronal microscale observatory arises from an intrinsic property of the Fresnel zone plate and its cousins: The focal length is so long at short wavelengths that the telescope must be distributed between two spacecraft separated by 0.1–1 km. With the optic on one and the detector on the other, the relative positions of the two spacecraft must be controlled to millimeter accuracy. The technology for such precision formation flying is developing rapidly. Thus, a coronal microscale observatory that implements a classical optical concept using state-of-the-art technology may solve a solar mystery that has endured for the better part of a century.

REFERENCES

1. E. W. Cliver, W. F. Dietrich, *J. Space Weather Space Clim.* **3**, A31 (2013).
2. R. Ramaty, R. G. Stone, eds., *High Energy Phenomena on the Sun*, NASA (1973).
3. J. A. Klimchuk, *Sol. Phys.* **234**, 41 (2006).
4. T. R. Rimmele et al., *Sol. Phys.* **295**, 172 (2020).
5. S. Tomczyk et al., *J. Geophys. Res. Space Phys.* **121**, 7470 (2016).
6. S. Gosain et al., in *Ground-Based and Airborne Instrumentation for Astronomy III*, C. J. Evans, L. Simard, H. Takami, eds., SPIE (2018), p. 107024H.
7. T. S. Bastian, in *Solar and Space Weather Radiophysics: Current Status and Future Developments*, D. E. Gary, C. U. Keller, eds., Springer (2004), p. 47.
8. D. E. Gary et al., https://arxiv.org/abs/2210.10827.
9. G. E. Brückner et al., *Sol. Phys.* **162**, 357 (1995).
10. D. Rabin, in *The WSPC Handbook of Astronomical Instrumentation*, vol. 3, A. M. Moore, ed., World Scientific (2021), p. 339.

WIKIMEDIA COMMONS/PUBLIC DOMAIN

Emmy Noether, pictured here circa 1900, was one of the first German women to receive a PhD in mathematics.

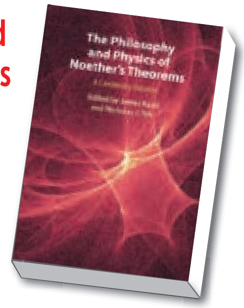
Celebrating Emmy Noether

The mathematician Emmy Noether (1882–1935) is best remembered by physicists for the systematic connection between symmetries and conservation laws, which she proved in her groundbreaking 1918 paper "Invariante Variationsprobleme" (Invariant variation problems). *The Philosophy and Physics of Noether's Theorems: A Centenary Volume*, edited by James Read and Nicholas Teh, is a celebration of Noether's 1918 paper and the ongoing importance of and fruitfulness of the theorems therein. It contains chapters by historians, physicists, and philosophers covering material ranging from biographical details to present-day applications of her theorems.

The decision to produce this collection is enormously welcome. In the decades after the publication of Noether's paper, it was rarely cited and even less frequently read. Indeed, physicists who made use of "Noether's theorem" were, until recently, often surprised to discover that her paper contained *two* theorems, both of which are important for the field.

An example of the more famous first theorem is the connection between rotational symmetry and conservation of angular momentum. The second theorem underlies important mathematical relationships, such as the Bianchi identities in general relativity. Work on Noether's theorems has blossomed over the past quarter century and, as the volume editors point out, includes an increasing number of applications across a wide variety of physics subfields. The paper's

The Philosophy and Physics of Noether's Theorems
A Centenary Volume
James Read and Nicholas J. Teh, eds.
Cambridge U. Press, 2022. $155.00



centenary saw several commemorative conferences, but the present volume is, to my knowledge, the only one dedicated to the 1918 paper.

The Philosophy and Physics of Noether's Theorems is not an introduction to her theorems, nor is it designed to be. It is a collection of papers presenting the state of the field. For that reason, readers with different backgrounds will find certain papers more—or less—attractive and accessible. Although it is unlikely that many people will read the book from cover to cover, anyone with an interest in Noether's theorems will find something to draw them in. Each contribution stands on its own, and the editors' introduction is a succinct and helpful guide to finding one's way around the 14 chapters.

The opening chapter, by Yvette Kosmann-Schwarzbach, provides an accessible overview of the context, content, and reception of Noether's 1918 paper, and it is nicely complemented by chapters 2 and 3, which are primarily historical. We learn about the great mathematicians Felix Klein and David Hilbert inviting Noether to join them at the University of Göttingen in 1915, shortly before Albert Einstein visited. She became involved in their investigations into the as-yet incomplete general theory of relativity, which led her to formulate her two theorems. The biographical details of Noether's 1933 dismissal from Göttingen by the Nazis and the efforts of many colleagues to find a safe place for her outside Germany—in the USSR, the UK, and, successfully, the US—remind us of the importance of our academic community and our support for one another in difficult times.

Together, the first three chapters introduce themes that run through multiple later contributions. One of those is energy conservation. It played a role in the puzzles that led Noether to formulate her theorems and continues to pose questions in the context of general relativity today. Readers are given multiple angles from which to see how Noether's theorems can be used to probe the mathematical structure of various theories and the ways in which energy conservation is encoded in that structure.

A second theme running throughout the book is the subtleties that arise once we move beyond the familiar—and simplified—derivations of Noether's theorems. The issues discussed include the relationship between variational symmetries and other types of symmetry, especially those to which we attribute physical significance; the converse theorems; how to state the most general formulation of the theorems; and whether symmetries are more fundamental than conservation laws. In chapter 7, Harvey Brown argues that the alleged primacy of symmetries over conservation laws is because of their heuristic power and pragmatic utility rather than any physical (or even metaphysical) priority. And that brings us to a third theme, which is the wide range of contexts in which Noether's theorems find useful application. Examples in the book include general relativity, classical particle mechanics, quantum electrodynamics, algebraic quantum field theory, the theory of defects in elastic media, and the heat equation.

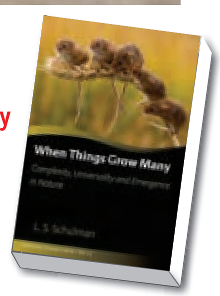
Because of how these and other themes recur, there are extra benefits for someone reading the volume from beginning to end. But readers who encounter a chapter that is not to their taste or for whom the mathematical demands outstrip their expertise should simply move on to the next chapter.

In 2003 Elena Castellani and I collected a set of papers on the general topic of symmetry in our edited volume *Symmetries in Physics: Philosophical Reflections*. Noether's theorems were just one theme among many discussed in our book. Almost two decades later, Read and Teh have produced a volume devoted entirely to Noether's theorems. It simultaneously serves as a comprehensive demonstration of the past 20 years of progress in philosophy of physics, an invaluable reference for physicists and philosophers alike, and a superb springboard for future research.

Katherine Brading
Duke University
Durham, North Carolina

How the justices of the US Supreme Court will vote on cases is an example of a complex system and one that can be predicted using the maximum-entropy method.

Exploring complex systems through applications



When Things Grow Many
Complexity, Universality and Emergence in Nature
L. S. Schulman
Oxford U. Press, 2022. $59.00

There is no universally accepted definition of a complex system, but it is often characterized as a composition of many interacting components that display emergent properties, such as self-organization, power-law distributions, and phase transitions. The complex-systems community is highly interdisciplinary, and the systems studied usually originate from disciplines outside of physics.

Because complex systems share some of the characteristics of condensed-matter systems—namely, many elementary components, multiple scales, and phase transitions—it is natural to apply statistical mechanics techniques to them. In his new textbook, *When Things Grow Many:*

Complexity, Universality and Emergence in Nature, Lawrence Schulman does just that. Beginning with an introduction to standard statistical mechanics techniques, he then explores their application to a curated set of examples from the literature on complex systems.

In the first half of the book, Schulman introduces readers to probabilistic techniques, the mean-field approximation and how it can fail because of fluctuations, bifurcations, stability analysis, critical phenomena, and master equations. He uses clearly explained examples from physics, including the ideal-gas law, ferromagnetism, and galaxy formation, and from other fields, such as epidemiology. Although those topics could occupy an entire statistical mechanics course, Schulman adeptly breezes through them while clearly highlighting their main concepts.

Schulman's inclusion of information theory and the maximum-entropy method, covered in chapter 6, is a welcome addition. To his credit, he focuses on a technique first developed by Edwin Jaynes that is not typically taught in physics curricula but deserves to be more widely known. In the 1950s Jaynes showed that the Boltzmann distribution can be derived by using information theory and the available information to find the most unbiased probability distribution. That approach treats the prediction of a system with many particles as a statistical inference problem, avoids thorny fundamental issues like ergodicity, and applies to fields well beyond physics. Schulman uses the Jaynes approach to derive the Maxwell velocity distribution and shows how it can be applied to develop a thermodynamic theory of ecosystems and even predict how US Supreme Court justices will vote.

In the second half of the book, Schulman applies statistical mechanics to a wide range of interesting topics, including traffic flow, flocking, galaxy morphology, segregation of urban neighborhoods, and synchronization. He introduces each problem well and provides entry points to the relevant literature.

One of the most challenging aspects of modeling complex systems is that their elements are often heterogeneous—in contrast to the identical atoms in condensed matter—and those elements' interactions are not necessarily short ranged and do not have the nearest-neighbor topology of atomic systems. I wish *When Things Grow Many* had more extensively covered techniques for tackling those issues, such as agent-based modeling and network theory, because easy-to-learn tools to do so are readily available.

I enjoyed reading *When Things Grow Many* and learned something new from each chapter. Schulman writes in a conversational style, and he peppers the book with jokes and opinions. Even though he intimates that he doesn't have all the answers, his fun, inviting tone will make readers want to find out if he does. Scattered throughout the book are many computational and analytical exercises, some of which are open ended. The book also contains an extensive set of appendices with brief reviews of useful topics like probability and stochastic dynamics. I expect that anyone interested in complex systems and who has the requisite knowledge of elementary calculus and linear algebra will find *When Things Grow Many* to be a rewarding read.

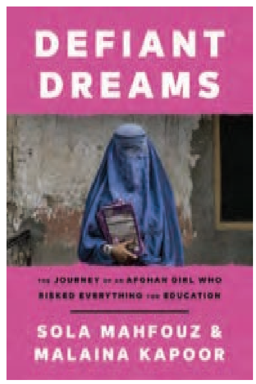
Robert Deegan
University of Michigan
Ann Arbor

NEW BOOKS & MEDIA

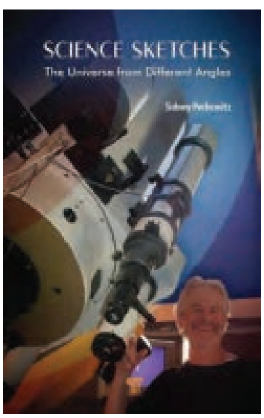
Defiant Dreams

The Journey of an Afghan Girl Who Risked Everything for Education

Sola Mahfouz and Malaina Kapoor
Ballantine Books, 2023. $28.00



It's now been two years since US-led forces withdrew from Afghanistan and the Taliban retook control of the country. This bracing new memoir by Sola Mahfouz, a pseudonymous Afghan refugee currently studying quantum computing at Tufts University, is a reminder of the stakes at play for everyday Afghan people. Born in 1996, Mahfouz lived through Taliban rule, the 2001 US-led invasion, and civil war before fleeing the country in 2016. Overcoming resistance toward women's education from both family members and the Taliban, she started studying English at home and eventually made it to the US with the help of an online conversation partner. *Defiant Dreams* is both an inspiring story and a tragic reminder of the human cost of war. —RD



Science Sketches

The Universe from Different Angles

Sidney Perkowitz
Jenny Stanford, 2022. $49.95

Science Sketches comprises 52 published articles and essays by Sidney Perkowitz, an emeritus physics professor and popular science writer. Similar to his first anthology, *Real Scientists Don't Wear Ties*, which appeared in 2019 (see Physics Today, March 2020, page 52), the material is written for general readers and falls roughly into three categories: science, technology, and culture. In addition to overviews of scientific concepts such as electromagnetic waves, black holes, and quantum gravity, Perkowitz discusses books, art, and films that present math and science concepts; biographies of scientists; and noteworthy technology and its roles in current affairs. Through his use of anecdotes, history, and nontechnical language, Perkowitz strives to present some fairly weighty scientific concepts to nonscientists. —CC PT

NEW PRODUCTS

Focus on software, data acquisition, and instrumentation

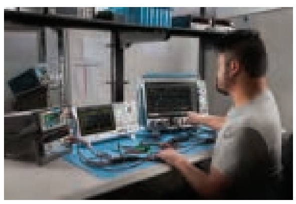
The descriptions of the new products listed in this section are based on information supplied to us by the manufacturers. Physics Today can assume no responsibility for their accuracy. For more information about a particular product, visit the website at the end of its description. Please send all new product submissions to ptpub@aip.org.

Andreas Mandelis

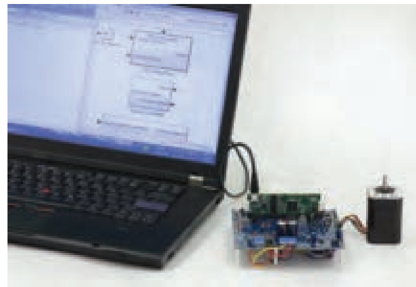
Polarity-reversing modules

Spellman High Voltage has launched its MXE family of high-voltage, hot-switchable, polarity-reversing modules with output voltages of 2.5–10 kV at 200 μA. The ultracompact MXE series is controlled by means of an analog interface provided via a standard 15-pin D-type connector. The units feature a differential-voltage program input for low-noise control and "enable" and "polarity-control" signal inputs compatible with transistor–transistor logic. According to the company, its advanced low-noise, high-stability technology provides the high quality and performance needed for precision applications, such as mass spectrometry, automatic test equipment evaluation, capillary electrophoresis, dual-ion surface analysis, electrostatic printing and lensing, and electrospinning. The products are compliant with the following European and UK environmental protection, health, and safety regulations: CE, RoHS, and UKCA. ***Spellman High Voltage Electronics Corp***, *475 Wireless Blvd, Hauppauge, NY 11788, www.spellmanhv.com*

Software for double-pulse testing



A new release of the oscilloscope-based wide-bandgap double-pulse test (WBG-DPT) solution from Tektronix speeds up validation time on silicon carbide and gallium nitride power converters. The WBG-DPT solution provides automated, repeatable, and accurate measurements on devices such as SiC and GaN MOSFETs. It integrates seamlessly into the measurement system of the Tektronix 4, 5, and 6 series mixed-signal oscilloscopes. According to the company, the WBG-DPT solution features several industry-first measurement capabilities, such as an automatic WBG deskew technique that eliminates the need for rewiring and can even be performed after double-pulse measurements have been taken. Plots of reverse-recovery timing now make it easy for users to see reverse-recovery details for multiple pulses overlaid on a single display. They can zoom in on the reverse-recovery region and even debug reverse-recovery parameters of the system. ***Tektronix Inc***, *14150 SW Karl Braun Dr, PO Box 500, Beaverton, OR 97077, www.tek.com*

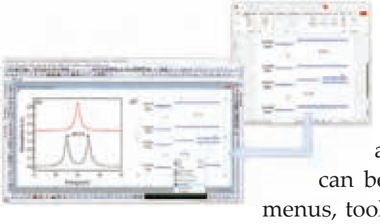


Mathematical programming software

Release 2023a (R2023a) of MathWorks MATLAB and Simulink software introduces MATLAB Test and enhanced features to help researchers and engineers develop, execute, verify, manage, and document dynamic tests of MATLAB code at scale. Organizations can group, save, and run custom test suites and identify untested code paths using industry-standard code-coverage metrics, such as condition, decision, and modified condition/decision coverage. The project-based quality dashboard can help users create an interactive, graphical summary of code-quality metrics and clickable details for code analysis and coverage, test results, and requirements. R2023a also features the new C2000 Microcontroller Blockset, with which users can design, simulate, and implement applications for Texas Instruments C2000 microcontrollers. They can also model digital power conversion and motor control applications. Major updates have been made to MATLAB and Simulink tools, including the Communications Toolbox and the Aerospace, Powertrain, and Vehicle Dynamics Blocksets. ***The MathWorks Inc***, *1 Apple Hill Dr, Natick, MA 01760-2098, www.mathworks.com*

Motion-control platform

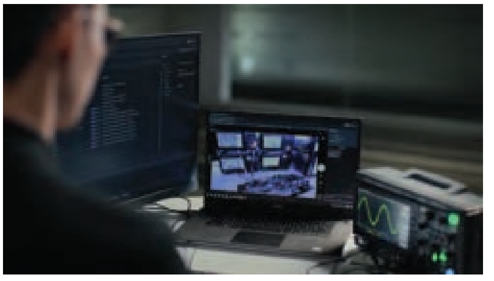
Aerotech has developed new products and features for its Automation1 precision-machine and motion-control platform. Two servomotor drives with integrated motion controllers can command 12 axes at 20 kHz over the glass optical fiber HyperWire motion bus. The Automation1-iXL2e, a compact single-axis linear-amplifier servo drive, offers high in-position stability and small-step-size operation. It generates very low electromagnetic interference noise. The Automation1-iXC6e pulse-width-modulation servomotor drive is designed for larger loads that require high power output. It includes a 480 V_{AC} (680 V_{DC} bus voltage) option. New features in the HMI (human–machine interface) Builder for Windows PCs MachineApps tool include enhanced layout flexibility with drag-and-drop controls, indicator lamps, and better state-machine control features. The Automation1 device catalog makes the advanced gantry-control configuration accessible to precision machine builders. They can add the gantry as a mechanical device that can be set up using the machine setup wizard. ***Aerotech Inc***, *101 Zeta Dr, Pittsburgh, PA 15238-2811, www.aerotech.com*



Data analysis and graphing software

Version 2023b of OriginLab's Origin and OriginPro data analysis and graphing software adds more than 150 new features, enhancements, and apps. To improve window management and let users take advantage of desktop space and multiple monitors, windows can be floated outside the Origin interface. Floating windows, with their own customizable menus, toolbar buttons, and Object Manager panel, can be pinned, dragged to an active folder, and retracted to their respective folder locations. Improvements to Browser Graph for multichannel data exploration include a new Stack Lines template, new controls to hide and sort plots, and faster run times. Built-in support for LaTeX equations has been added. New graph types in version 2023b include Sankey Map, Stacked Bar Map, Double-Y Waterfall, and Sunburst Color by Level. Apps let Origin users customize their software by allowing them to choose which tools they want. Those available for 2023 include Brain Atlas Analyzer, Abbott–Firestone Curve, Gaussian Filter in Surface Metrology, and Flow Cytometry Standard Connector. ***OriginLab Corporation***, *One Roundhouse Plaza, Ste 303, Northampton, MA 01060, www.originlab.com*

Digital learning platform

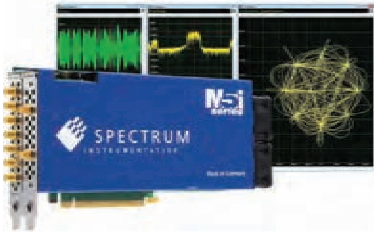


The Digital Learning Suite from Keysight offers university engineering educators and students one-stop access to laboratory resources, measurement data analysis tools, and learning resources relevant to industry through a single, secure Web interface. The platform streamlines and simplifies laboratory management and allows educators to focus on teaching and collaboration tools to interact with students in real time. The software suite integrates with learning management systems (LMSs), such as Blackboard and Canvas. It provides direct access to the soft front panel of laboratory instruments so users can perform measurements remotely, take screenshots, plot measurement data and math functions on an x–y graph, and more. The platform connects and links to instruments in a laboratory and on a local network. It provides access to laboratory exercises and courseware from Keysight University through a single sign-on from an integrated LMS. ***Keysight Technologies Inc***, *1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*

Antivibration cryostat

ICE Oxford has upgraded its Dry Ice 300 mK single-shot system. The new Tertia cryostat offers single-shot and continuous operation, an extended hold time of 100 h with 7 L of helium-3, a reduced recondense time of 1 h, and up to 10 ports for wiring. Users of the closed-loop system can carry out experiments in a 300 mm sample plate within a temperature range of 280 mK to 425 K. The system offers a high-cooling-power capability of 50 µW at 285 mK and 100 µW at 295 K. It can achieve ±1 mK and maintain temperature stability at 300 mK and above; its cool-down time to base temperature is 22.5 h. Tertia uses the company's ICE sock technology: The cryocooler is supported in a cold antivibration frame "sock" that separates the system's two cold heads from the cryostat. Vibration levels at the sample space are reduced to about 20 nm; additional options can reduce them to less than 10 nm. The ability to counteract high vibration means that Tertia can take advantage of Gifford–McMahon cryocooler technology, which is less costly than pulse-tube refrigerator cooling. The ICE sock technology also allows for easy removal of the cold head and eliminates the need for a remote motor. ***ICE Oxford***, *Ave Four, Station Lane, Witney, Oxford OX28 4BN, UK, www.iceoxford.com*

PCIe digitizer cards



Spectrum Instrumentation has produced two new PCIe (peripheral component interconnect express) digitizer cards. The M5i.3350-x16 and M5i.3357-x16 extend the company's M5i series to deliver optimal signal acquisition and analysis capabilities to users working with challenging gigahertz-range electronic signals. The M5i.33xx series offers five models with one or two channels; users can select sample rates of 3.2, 6.4, or 10 GS/s and bandwidths of 1, 2, or 3 GHz. The cards combine an ultrafast sample speed of 10 GS/s with a vertical resolution of 12 bits and a data streaming rate of 12.8 GB/s. Sixteen-lane, Gen3, PCIe technology is used to achieve the high streaming rate that allows the acquired data to be sent directly to a computer's memory for storage or to central processing units and graphics processing units for customized signal processing and analysis. The cards feature front-end circuitry with a bandwidth of more than 3 GHz and onboard memory of up to 16 GB (8 GS). They can handle various signals, such as those found in communications, semiconductor testing, spectroscopy, optical systems, and quantum physics. ***Spectrum Instrumentation Corp***, *401 Hackensack Ave, 4th Fl, Hackensack, NJ 07601, https://spectrum-instrumentation.com* PT

Bo Miao is a postdoctoral associate, **Jaron Shrock** is a graduate student, and **Howard Milchberg** is a professor of physics and electrical and computer engineering at the University of Maryland in College Park. They are all members of the Intense Laser Matter Interactions group, of which Milchberg is the principal investigator.



Electrons see the guiding light

Bo Miao, Jaron Shrock, and Howard Milchberg

To accelerate electrons to multi-GeV energies with lasers, keep the bright light tight.

The electric field of a tightly focused milliwatt laser pointer can reach an impressive 50 000 V/m, a gradient that could, in principle, accelerate the electrons in a dental x-ray tube. With lots of cheap laser pointers and some assembly time at your kitchen table, can you build a 10 million V/m compact electron accelerator? Sadly, it wouldn't work, and you'd needlessly annoy the neighbors.

Three main problems undermine the laser-accelerator project. First, although the light wave in a laser beam is a coherent electromagnetic field oscillation, with peaks and valleys aligned in lockstep, the train of peaks and valleys from the second laser is randomly phase shifted in time with respect to the first. So the fields from multiple lasers would interfere, as peaks from one laser cancel valleys from another.

If one has N such randomly phased laser beams, each with intensity I, the peak intensity from combining the beams would be NI. If all of them were in phase, with their wave trains aligned, however, the peak intensity would be N^2I. But, alas, that dividend of coherent superposition is unavailable to the well-meaning hobbyist who just purchased a wheelbarrow full of $5 laser pointers.

Suppose you bypass the problem by simply buying a laser that, when tightly focused, gives an electric field of 10^7 V/m. Then the second problem can be expressed this way: "Hey genius, where's my meter?" The heckler is pointing out that the focused laser beam is, at most, tens of microns wide—far less than 1 m. So the field is better expressed as the ratio 200 V/(20 μm), and your laser purchase would provide (at most) only 200 V of accelerating potential across the focal spot. Actually, it wouldn't even remotely do that, courtesy of the third problem.

That problem is that laser light in free space is a high-frequency transverse electromagnetic wave, with fields orthogonal to the beam direction: An electron would do a high-frequency sideways shimmy, with negligible net acceleration and energy gain.

Mode locking

To solve the first problem (insufficient laser intensity), combining all the lasers together in phase would be a tough task. Fortunately, a method already exists that does the job in a single laser: It's called mode locking, and it allows many beams, or modes, in the laser to oscillate simultaneously. Although a mode-locked oscillator will cost orders of magnitude more than a pile of laser pointers, the mode-locking process produces a train of short, nanojoule-energy pulses whose peak intensity is N^2 as high as one of the individual modes, where N (here the number of modes) can be a very large number. The electric field in just one of the pulses, when focused, is well beyond 10 million V/m.

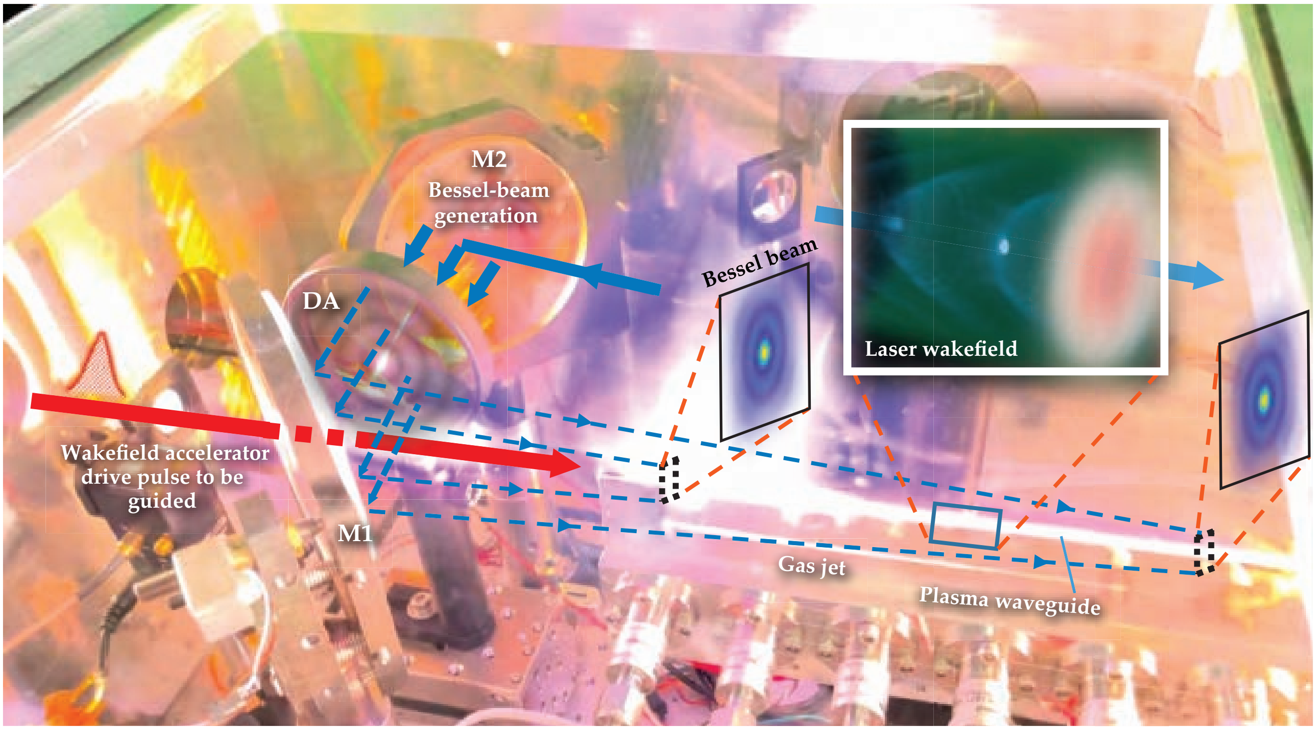
It turns out that even that large a field is much too weak to compete with conventional particle accelerators. With a lot more money for lasers and optics, however, and the important Nobel Prize–winning technique of chirped pulse amplification (see Physics Today, December 2018, page 18), one of those nanojoule pulses can be amplified by a factor 10^{10} to get a 10 J laser pulse of duration 50 fs. Focusing that pulse to, say, a 50 μm spot then gives intensities greater than 10^{19} W/cm^2, which corresponds to a peak electric field of 5×10^{12} V/m. That's an impressive 0.25 GeV across the focal spot. Electron acceleration modules at the GeV level are of interest for x-ray light sources and can form stages for future high-energy-physics applications.

Plasma waveguides to the rescue

But don't get too excited (yet). The sideways shimmy would ensure that a negligible fraction of the 0.25 GeV would be imparted to an electron. But in a neat trick of plasma physics and optics, the second and third problems—how to get a long acceleration distance and how to beat the high-frequency transverse field—are both solved using an optical fiber made of plasma. As discussed in one of the authors' earlier works—see the Quick Study by Howard Milchberg, Physics Today, June 2019, page 70—such a fiber, called a plasma waveguide, can optically guide an ultra-intense laser pulse over long distances.

A plasma waveguide can easily support peak intensities of at least 10^{19} W/cm^2. That's 10^8 times as high as the damage threshold of a conventional glass optical fiber. As the intense pulse propagates down the plasma waveguide, its pressure pushes the lighter electrons out of the way—leaving the heavier protons unperturbed—and excites a plasma charge-density wave that follows the pulse like a water wake from a boat, as shown in the figure. That charge-density wake supports an enormous electrostatic field (the so-called laser wakefield) that points along the pulse-propagation direction. (See the article by Wim Leemans and Eric Esarey, Physics Today, March 2009, page 44.)

In our recent experiments using the ALEPH (Advanced Laser for Extreme Photonics) at Colorado State University, a 25 GV/m laser wakefield was induced by a 10^{19} W/cm^2 laser pulse injected into and optically guided by a 20-cm-long hydrogen plasma waveguide narrower than a human hair. The electrons were injected into the wakefield from laser ionization of a small amount of nitrogen added to the hydrogen gas and accelerated.



A 280 TW LASER PULSE (red arrow) is injected through a hole in mirror M1 and into a 20-cm-long hydrogen-gas density depression generated a few nanoseconds earlier from gas heating by a copropagating Bessel beam. The leading portion of the drive pulse forms its own plasma waveguide as it propagates along the density depression, simultaneously exciting a wakefield in the plasma (the top right inset shows a simulation of the laser pulse and its trailing wakefield). The Bessel beam is generated by a separate synchronized laser pulse (blue arrows) that passes through a diffractive axicon DA (between mirrors M2 and M1) and focuses into the hydrogen-gas jet. The insets show the measured intensity profiles of the Bessel beam. (Background photo courtesy of Reed Hollinger, Colorado State University.)

The plasma is effectively a field transformer: Although electron acceleration to 0.25 GeV across the focal spot of a 10^{19} W/cm^2 electromagnetic pulse is unattainable in free space, sending that pulse down a long plasma waveguide drives an electrostatic acceleration 20 times as high—up to 5 GeV. (It's worth noting that electron beams from conventional high-energy linear electron accelerators can also drive plasma wakefield acceleration and produce energy-doubled beams.)

A plasma waveguide obeys the same principle as all optical fibers: The refractive index near the axis, or core, of the waveguide must be higher on axis than off axis in the fiber's periphery—the cladding. In a plasma, where the polarizability of free electrons is negative, that means the core must have lower density than the cladding. We recently demonstrated two methods for generating such extended annular plasmas using Bessel-beam laser pulses.

In the scheme used in our acceleration experiments—pictured here—an ultrashort laser pulse with a transverse intensity profile shaped like a zeroth-order Bessel function (a J_0 pulse) first ionizes hydrogen gas from a long gas jet, forming a tens-of-centimeters long and very thin (less than a 10 μm width) plasma of temperature 10^5 K. The high-pressure plasma explosively expands in a few nanoseconds, snowplowing a long cylindrical shell of enhanced neutral hydrogen density—a shock wave—on the outside, with the on-axis density rapidly dropping.

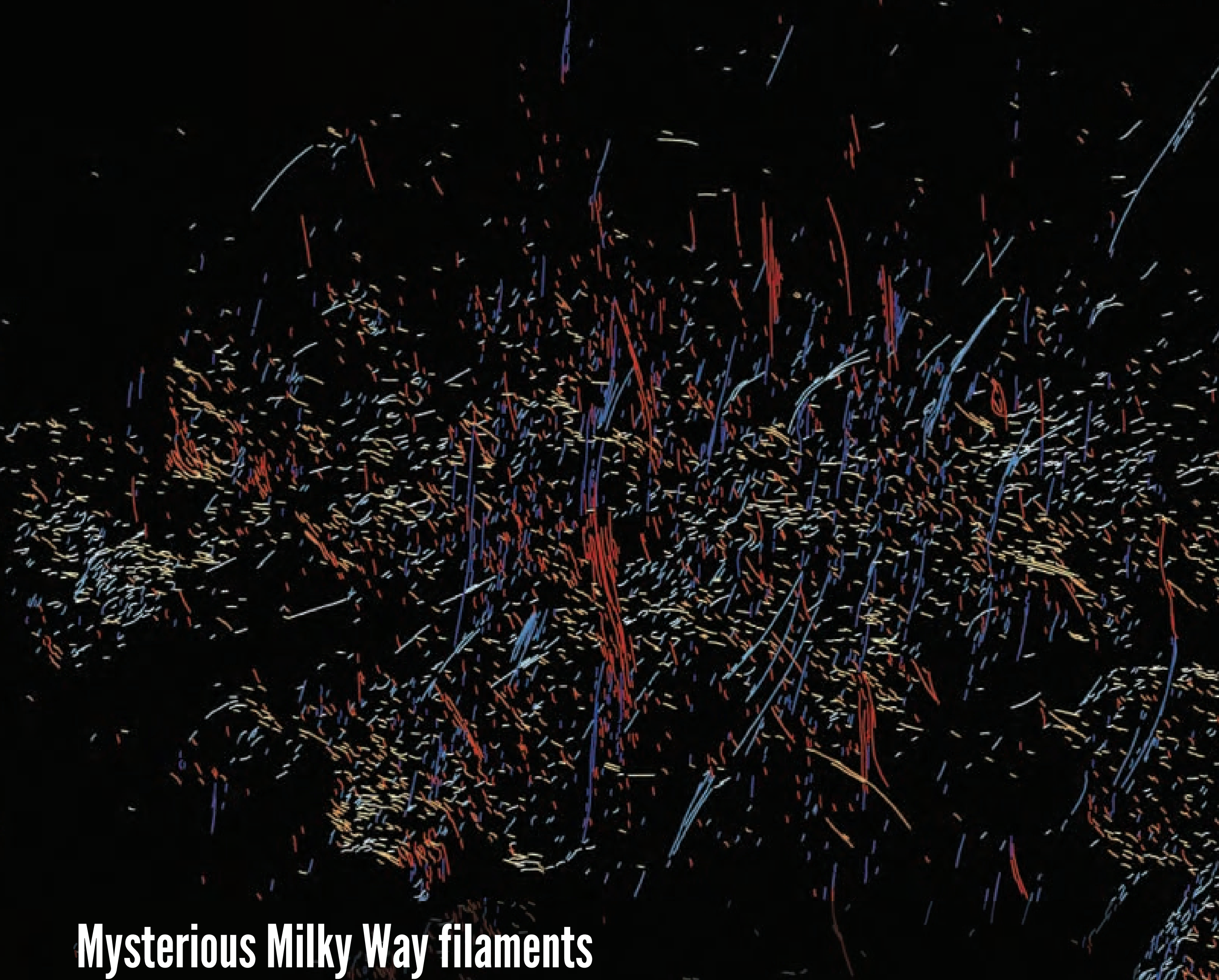
After a few nanoseconds, the long hydrogen gas structure formed by the low-density core and the surrounding neutral shell is injected at one end by the ultrahigh-intensity pulse. And then something initially unexpected and remarkable takes place: The leading portion of the injected pulse ionizes the core and the inside wall of the neutral shell, forming a plasma waveguide that guides the rest of the pulse—analogous to a locomotive laying railroad tracks for the rest of the train cars to follow. Not only does that self-waveguiding pulse generate its own plasma waveguide as it propagates, but it also excites the laser wakefield responsible for our measured acceleration up to 5 GeV in 20 cm.

So what advice would we give a young scientist interested in advanced accelerator physics? One word: Plasmas. They're the accelerator medium of the future.

Additional resources

► T. Tajima, J. M. Dawson, "Laser electron accelerator," *Phys. Rev. Lett.* **43**, 267 (1979).
► E. Esarey, C. B. Schroeder, W. P. Leemans, "Physics of laser-driven plasma-based electron accelerators," *Rev. Mod. Phys.* **81**, 1229 (2009).
► B. Miao et al., "Optical guiding in meter-scale plasma waveguides," *Phys. Rev. Lett.* **125**, 074801 (2020).
► L. Feder et al., "Self-waveguiding of relativistic laser pulses in neutral gas channels," *Phys. Rev. Res.* **2**, 043173 (2020).
► B. Miao et al., "Multi-GeV electron bunches from an all-optical laser wakefield accelerator," *Phys. Rev. X* **12**, 031038 (2022).
► I. Blumenfeld et al., "Energy doubling of 42 GeV electrons in a metre-scale plasma wakefield accelerator," *Nature* **445**, 741 (2007).

PT



Mysterious Milky Way filaments

Around 25 000 light-years away, high-energy particles are moving near the speed of light in large vertical magnetized threads perpendicular to the plane of the Milky Way. The particles are likely still zipping around inside them today. Northwestern University's Farhad Yusef-Zadeh and his colleagues first discovered the filament-like structures near Sagittarius A* (Sgr A*), the black hole at the center of our galaxy, almost 40 years ago. As of 2022, about 1000 filaments stretching roughly 150 light-years have been counted. Now, through a MeerKAT radio telescope survey of the galactic center, Yusef-Zadeh and other researchers have found, to their surprise, what they suspect to be a few hundred horizontal filaments 5–10 light-years in length that are pointing radially toward Sgr A* and parallel to the galactic plane.

To make this image, Yusef-Zadeh and his colleagues filtered the original MeerKAT image to smooth the background noise, and then they applied an algorithm technique to identify and quantify every filament, each of which is represented by a dash. The filaments here span a 3.5° x 2.5° field of view and display a colorful glimpse of only the inner few hundred parsecs of the Milky Way. The redder a filament is, the closer it points to the galactic north, whereas the bluer filaments point farther away. Yusef-Zadeh and his colleagues say that the vertical filaments do not have a clear energy source, but they suspect that the horizontal ones stem from jet-driven outflow from Sgr A*. Although they have no clear answers as to what the purpose is of both filament types, the scientists say that the horizontal filaments help further the understanding of Sgr A* and its accretion disk orientation. (F. Yusef-Zadeh et al., *Astrophys. J. Lett.* **949**, L31, 2023; image courtesy of Farhad Yusef-Zadeh.) —HHM

TO SUBMIT CANDIDATE IMAGES FOR BACK SCATTER VISIT https://contact.physicstoday.org.