PHYSICS TODAY

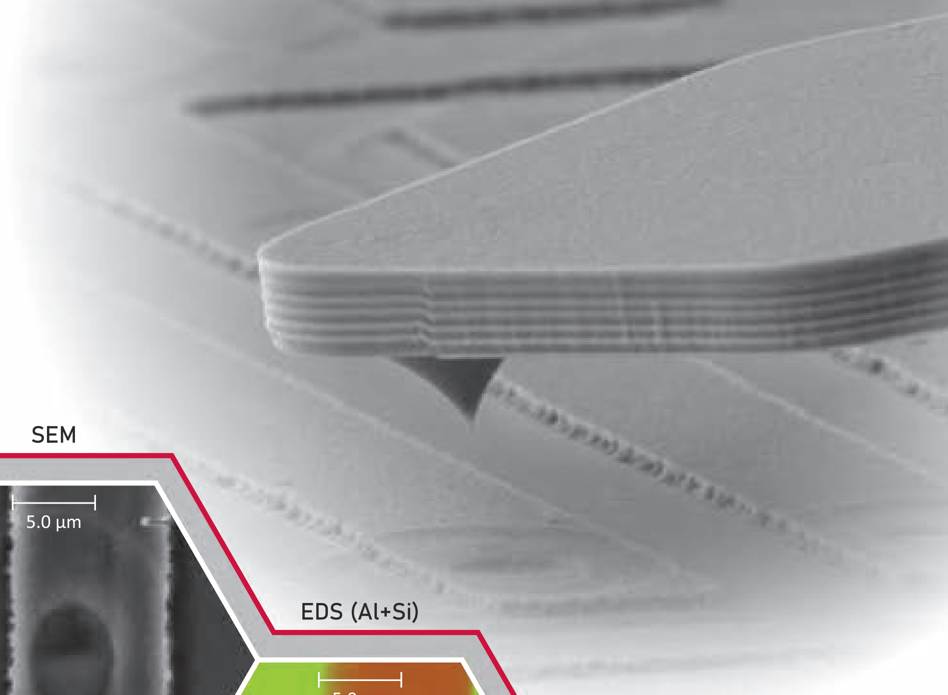
July 2024 • volume 77, number 7
A publication of the American Institute of Physics

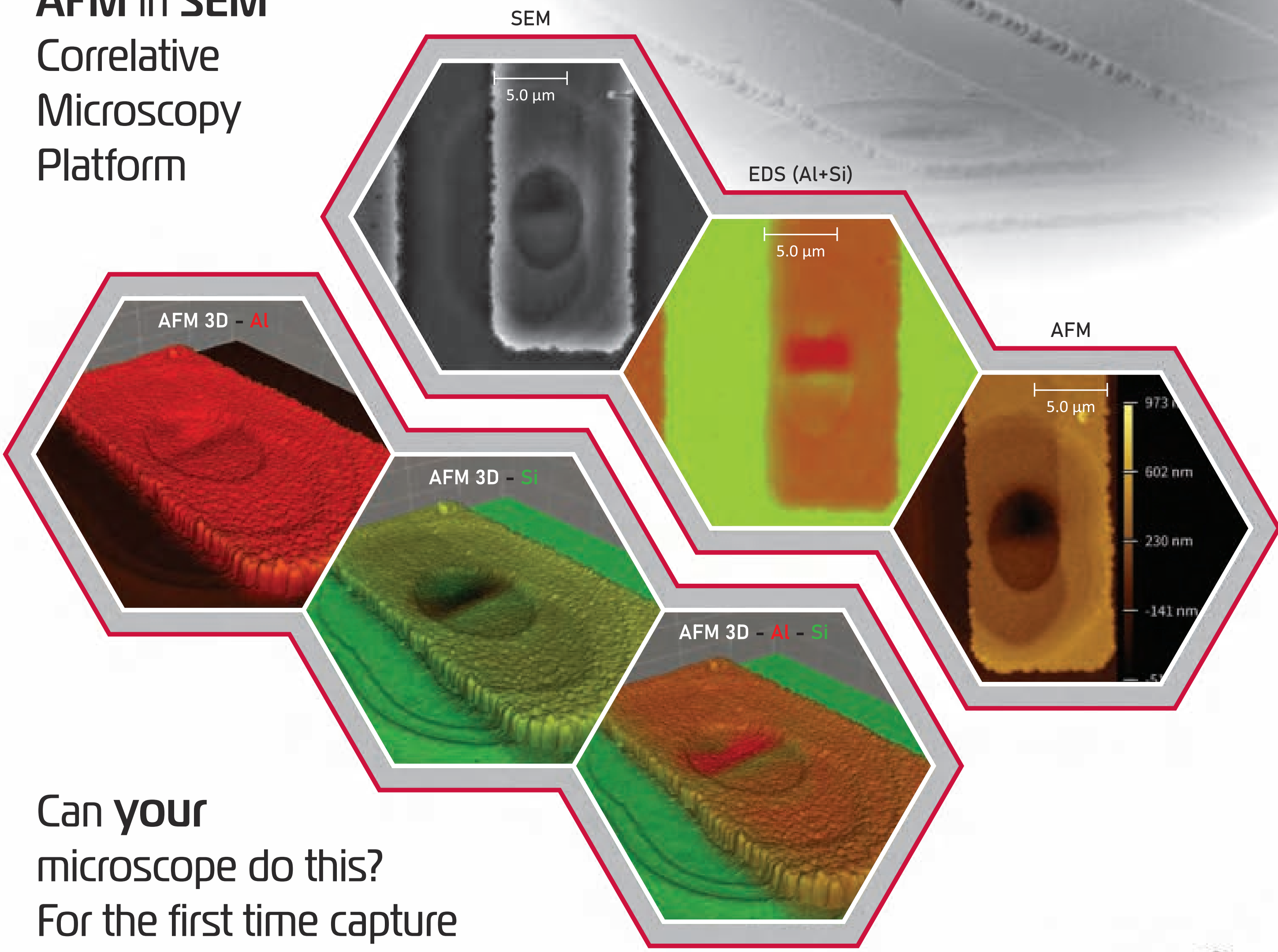
IMAGING POLARIZED LIGHT

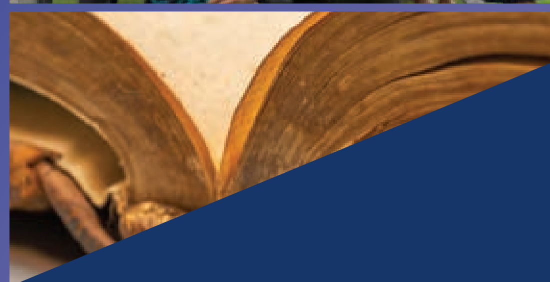
Exploring Mars's harsh atmosphere

Fixing the PhD qualifying exam

The threat from cosmic flotsam



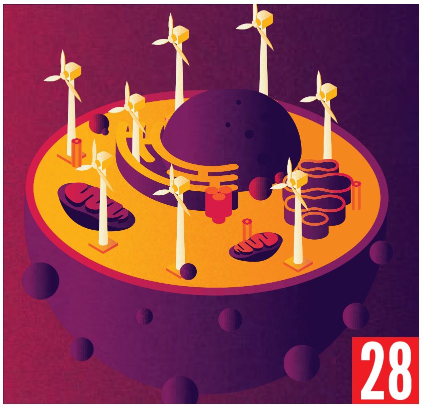

SEM
5.0 µm
EDS (Al+Si)
5.0 µm
AFM
5.0 µm
602 nm
230 nm
-141 nm
AFM 3D - Al
AFM 3D - Si
AFM 3D - Al - Si

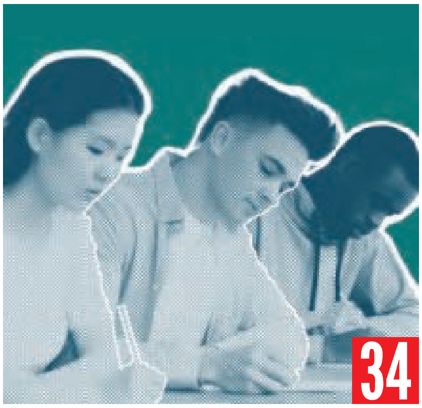


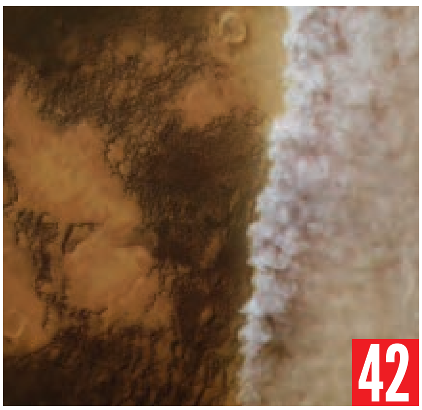
PHYSICS TODAY

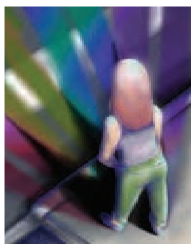
July 2024 | volume 77 number 7

FEATURES





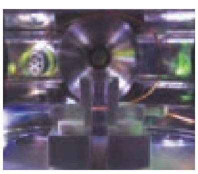




ON THE COVER: The light all around us comes in a multitude of polarization states. By splitting an image into four components, represented here by the colored shadows, a specialized camera can capture them all. Imaging the ways that objects can transform the polarization of the light that hits them is a task four times as complicated, but that can now be done with a compact device based on optical metamaterials. To learn more, turn to **page 12.** (Image by Freddie Pagani.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org



YALE WRIGHT LAB

Nuclear-recoil sensing

A new technique for studying radioactive decays replaces complicated particle detectors with tiny optically trapped spheres. Each time a radioactive nucleus in a sphere decays, the nucleus recoils, crashes into a neighboring atom, and causes an observable shift in the sphere's position.

physicstoday.org/Jul2024a

ALEXANDER POLLAK/U. CHICAGO

NSF delays CMB-S4

A planned $800 million project to study ancient cosmic radiation using telescopes at the South Pole and in Chile is getting delayed because of the need to upgrade infrastructure in Antarctica. The recent NSF decision has implications for CMB-S4 scientists and others doing research near the South Pole.

physicstoday.org/Jul2024b

WILL LAB

Dipolar-molecule BEC

Nearly three decades after the development of the first Bose–Einstein condensate, researchers have created a BEC composed of dipolar molecules. Such systems could enable scientists to study quantum magnetism and exotic phases of quantum matter, such as supersolids.

physicstoday.org/Jul2024c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 77, number 7. 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

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

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

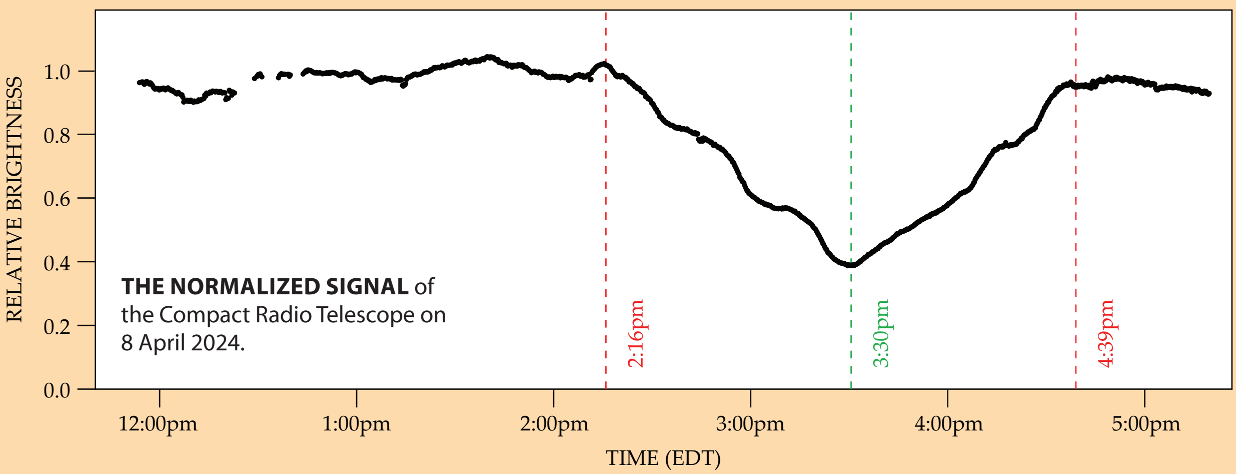
PhysicsToday **@physicstoday**

Live streaming a radio-telescope observation of the solar eclipse

Millions of people were in the path of totality of the Great North American Eclipse on 8 April. They had a rare opportunity to witness the corona, which is visible only during a total solar eclipse, as explained in Jay Pasachoff's August 2019 Quick Study (page 66). At Northeastern University in Boston, the eclipse provided us with a unique outreach opportunity involving our new Compact Radio Telescope (CRT), which is mainly used for 21 cm hydrogen-line observations in undergraduate and graduate physics teaching laboratories.

Starting a couple of hours before the eclipse, we tracked the Sun with the CRT and periodically measured the total power received in a bandwidth of a few megahertz around 1.42 GHz, the frequency of the 21 cm hyperfine transition in neutral hydrogen. The signal started to decrease close to the moment of optical Sun–Moon contact and reached a minimum of about 40% of what was measured before and after the eclipse. The visible solar photosphere occultation in Boston was 93%. We observed a smaller dip because additional radio emission comes from the vast solar co-

THE COMPACT RADIO TELESCOPE, on the roof of Churchill Hall at Northeastern University in Boston.



THE NORMALIZED SIGNAL of the Compact Radio Telescope on 8 April 2024.

rona,[1] which is normally concealed from the human eye by the blue sky.

The 2.5-hour solar eclipse gave plenty of time for us to engage and motivate students to learn more about astrophysics and radio astronomy. Many Northeastern students who watched our live stream[2] were curious about the difference between the optical and radio observations. The students were quite enthusiastic about the real-time radio tracking of the eclipse, so we plan to continue live streaming radio-telescope observations in the future.

References

1. V. V. Zheleznyakov, *Radio Emission of the Sun and Planets*, J. S. Hey, ed., H. S. H. Massey, trans., Pergamon Press (1970).
2. Northeastern University Center for Radio Astronomy, "Solar Eclipse 2024," https://cfra.sites.northeastern.edu/solar-eclipse-2024.

Baris Altunkaynak
(i.altunkaynak@northeastern.edu)
Oleg Batishchev
(o.batishchev@northeastern.edu)
Northeastern University
Boston, Massachusetts

Beyond the cinematic feat: Consequences *Oppenheimer* ignored

A big red button. A bright flash of light. A loud thundering sound. Laughter. Cheers. That is how Christopher Nolan's *Oppenheimer* depicts the 1945 Trinity nuclear test. The movie won seven Oscars, including Best Picture, as well as acclaim at other awards ceremonies. But as physicists and nuclear security researchers, we contend that the deleterious consequences and pervasive neglect that characterize the Manhattan Project—and the nuclear enterprise in general—merit as big of a spotlight as the film has received.

While the film glorifies the Trinity explosion as a scientific feat, it fails to fully convey its consequences. Although just one scene in the film, it is representative of how the movie misleads the public about the nuclear enterprise and romanticizes its scientific pursuits.

Despite the celebratory nature with which Trinity is portrayed in the movie, the test was in fact a tragedy. The project scientists were aware of the hazards posed by radiation,[1] yet the film barely recognizes the far-reaching dangers of fallout. The test led to contamination of air, food, and water sources in New Mexico. Tina Cordova, a cofounder of the Tularosa Basin Downwinders Consortium, has described how her family is one of the many New Mexican families who have experienced four or five generations of cancer.[2] Simulations by Sébastien Philippe of Princeton University and colleagues indicate that Trinity radiation made its way to 46 US states, Mexico, and Canada within 10 days of the explosion.[3]

Oppenheimer dangerously seems to celebrate scientists' role in contributing to one of the most dangerous scientific projects in human history. Other research by Philippe, for example, shows that a nuclear attack on central-US missile silos could potentially result in millions of deaths across North America.[4] Models by Rutgers University climate scientists Lili Xia and Alan Robock and their collaborators have shown that immediate climate changes caused by a nuclear war between Russia and the US could result in more than 5 billion deaths from starvation alone within two years.[5] Scientists, nevertheless, have continued working to maintain and enhance the US's collection of such weapons.

The movie, furthermore, underplays the agency of scientists by indicating a separation between the scientific and military components of the Manhattan Project. Perhaps the most prominent depiction of that separation is captured in the scene in which military officials arrive at Los Alamos to take away the bombs. When J. Robert Oppenheimer attempts to inform one official that the blast would be less powerful if the weapon is detonated high in the air, the man responds, "With respect, Dr. Oppenheimer, we'll take it from here." The scene conveys a message that although the scientists invented and constructed the atomic bombs, they were not responsible for the manner in which they were used. It's a depiction that could lead the public to absolve the scientists who actively participated not only in the bomb's construction but also in its deployment.

We can't argue with the movie's artistic quality. When you have access to such an incredibly wide-reaching platform, however, you have an enormous responsibility to get things right. The public, after all, may take the information in the film at face value. When you're conveying a message about nuclear weapons, it's important not to underplay the risks or the agency of scientists, even if unintentionally. At the end of the day, future scientists may be watching. We don't want to risk the chance that the next generation of scientists will continue to support the development of even-deadlier weapons and will work under the rationale that their projects are for the sake of science without fully considering the risks.

By romanticizing the development of nuclear weapons, *Oppenheimer*, like its namesake, might have done a disservice to society—potentially recruiting to the scientific enterprise more people who believe that science can be separated from its military use.

References

1. US Department of Energy Office of History and Heritage Resources, "Safety and the Trinity Test," https://www.osti.gov/opennet/manhattan-project-history/Events/1945/trinity_safety.htm.
2. T. Cordova, "What 'Oppenheimer' doesn't tell you about the Trinity test," *New York Times*, 30 July 2023.
3. L. M. M. Blume, "Trinity nuclear test's fallout reached 46 states, Canada and Mexico, study finds," *New York Times*, 20 July 2023; S. Philippe et al., https://arxiv.org/abs/2307.11040.
4. S. Philippe, "Who would take the brunt of an attack on U.S. nuclear missile silos?," *Scientific American*, 1 December 2023.
5. L. Xia et al., *Nat. Food* **3**, 586 (2022).

Bárbara Cruvinel Santiago
(barbcs@stanford.edu)
Lindsay Rand
(randl@stanford.edu)
Stanford University
Stanford, California PT

Metamaterial device makes 16 polarization measurements at once

Capturing all the ways that an object can affect a light wave's polarization has always been cumbersome. Now it can be done in an instant.

When you're walking on the beach on a sunny day, polarized sunglasses come in handy: The glare off the water, even though it's a reflection of unpolarized sunlight, is horizontally polarized, so it's blocked by the vertically polarized lenses. The phenomenon is the result of how light reflects off a dielectric, so glare off metal surfaces is different. If you had full polarization vision, you could readily tell whether a shiny object in the distance is a pool of water or a sheet of steel.

What would "full polarization vision" entail? To completely describe light's polarization state—whether linear, circular, elliptical, or partially or completely unpolarized—it takes four numbers, often represented as the four-component Stokes vector, named after George Gabriel Stokes, who worked out the idea in the 1850s. To get a fully polarized view of a scene, therefore, you'd need in effect to image it four times.

But that fourfold Stokes imaging doesn't capture all the polarization-changing properties of the objects you're looking at. It describes the polarization of the light coming out of the scene, but it says nothing about the light going in. A quarter-wave plate, for example, looks like an ordinary clear material under unpolarized light, but it transforms circularly polarized light into linearly polarized light and vice versa. To account for the possibility that any polarization state might be transformed into any other, you'd need a 4×4 matrix of images, or 16 in total.

You might think that imaging an object 16 times—a so-called Mueller matrix of images, after Hans Mueller's work in the 1940s—would require an intricate system of polarizers and wave plates to be carefully aligned and realigned. And until recently, you'd be right. But now Aun Zaidi (a recent graduate of Harvard University), his PhD adviser Federico Capasso, and colleagues have developed an imaging system that can do the job in a single optical shot.[1]

Instead of conventional optical components, the Harvard researchers used optical metamaterials—intricate arrays of tiny dielectric pillars—to control and analyze complex polarization states of light. The metamaterials are so compact, at just a few millimeters in diameter, that the whole imaging system can be packaged into a device that can go almost anywhere. It can even fit into an endoscope, where it would be ideally suited for one of polarization imaging's most

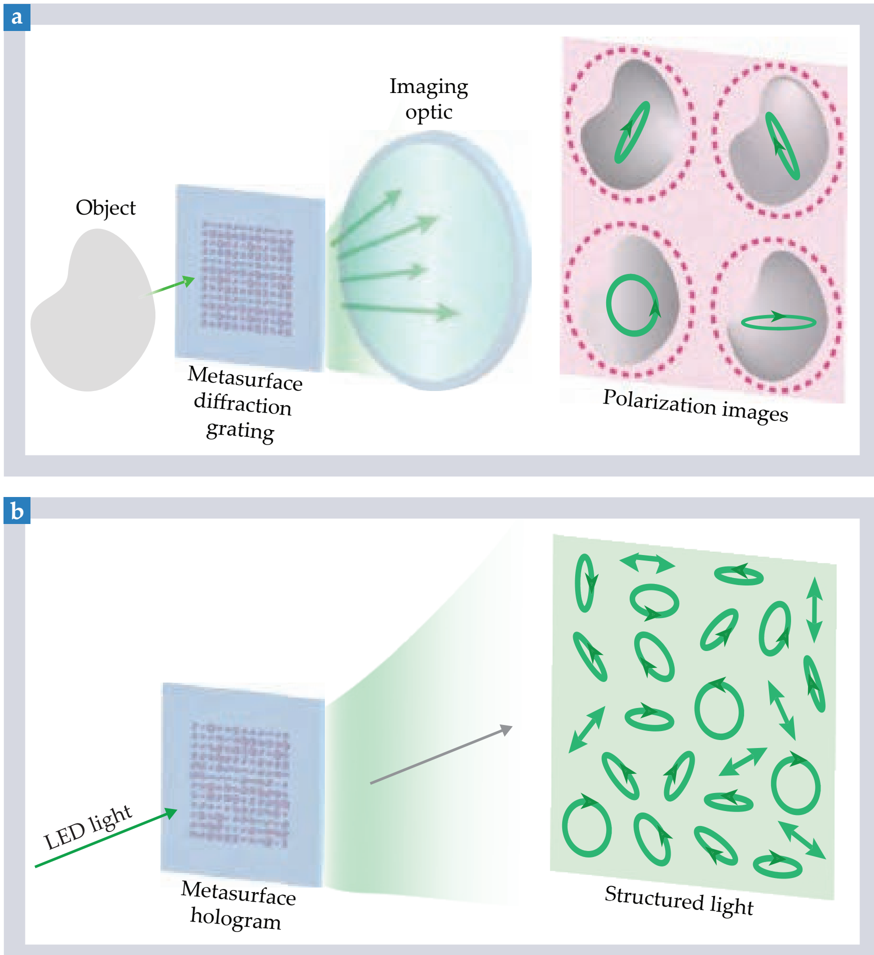


FIGURE 1. METASURFACE OPTICS made of tiny dielectric pillars have extraordinary power to shape the intensity and polarization patterns of light waves. **(a)** A metasurface diffraction grating can split the light coming off an object into four images of independent polarization states. **(b)** And a metasurface hologram can transform initially unpolarized light from an LED into structured light whose polarization varies on a fine spatial scale. Combining the two yields a system that can image an object's 4×4 Mueller matrix, which captures all of its polarization-changing properties. (Adapted from ref. 1.)

powerful applications: cancer screening.

Constructive interference

Ordinary materials get their properties—optical index of refraction, thermal conductivity, mechanical stiffness, and so on—from the interactions of atoms and molecules on the subnanometer scale. For metamaterials, in contrast, the salient features are orders of magnitude larger, but still small. The metasurface optics that the Capasso group uses are made of titanium dioxide pillars a few hundred nanometers wide, just smaller than the wavelength of visible light. Each one acts like a tiny waveguide that briefly slows down the light that passes through it. By controlling the pillar sizes and shapes, researchers can create interference patterns that both mimic and extend the effects of conventional optics.

The principle is easy enough to state, but the implementation is much more complicated. "If you want to create a metalens, that can be done beautifully," says Capasso. "You just arrange the pillars so all the light rays interfere constructively at the focus." (See, for example, PHYSICS TODAY, October 2022, page 19.) For anything more complicated than that, researchers quickly find themselves at the mercy of numerical methods: starting with a best guess, then iteratively refining it until the optical output is as close as it needs to be to the desired goal.

"It's hard to get an intuitive feeling for what the result means," says Capasso. "It's very frustrating, because as physicists we want to understand what we've done. But the algorithm is very powerful, and it spits out the solution."

In 2019 Capasso and colleagues, led by then-student Noah Rubin, presented a new mathematical formalism that allowed them to design metamaterials that controlled not just the light's intensity but also its polarization.[2] They then designed a polarization analyzer, shown schematically in figure 1a, that splits incoming light into four independent polarization components. The decomposition is different from the one Stokes used, but it's mathematically equivalent to imaging the Stokes vector.

Metasurface-based Stokes imaging may already be on its way to practical commercial use. The technology for fabricating optical metasurfaces is similar to what's already used to make semiconductor chips, so it's straightforward to integrate metasurfaces into phones, tablets, and other consumer electronics. And Capasso's startup company, Metalenz, is pursuing an application in enhancing the security of facial authentication.

Current facial-recognition tools are not foolproof: A thief could unlock a stolen phone by showing it a photo of the rightful owner. But photographs don't carry polarization information—and light reflected off skin is slightly polarized for much the same reason glare off the ocean is. Including a polarization analyzer in a phone's facial recognition camera can help ensure that it's unlocked only by the owner's actual face.

Four by four

As powerful as Stokes imaging is, it's just one step on the path to full 16-fold Mueller-matrix imaging. To measure an object's Mueller matrix, you'd need to probe its response to four different polarization states of light. You could do that sequentially in time, but that would require moving parts to physically swap out one metasurface element for another. Instead, Zaidi decided to try using structured light, whose polarization varies in space, as shown schematically in figure 1b.

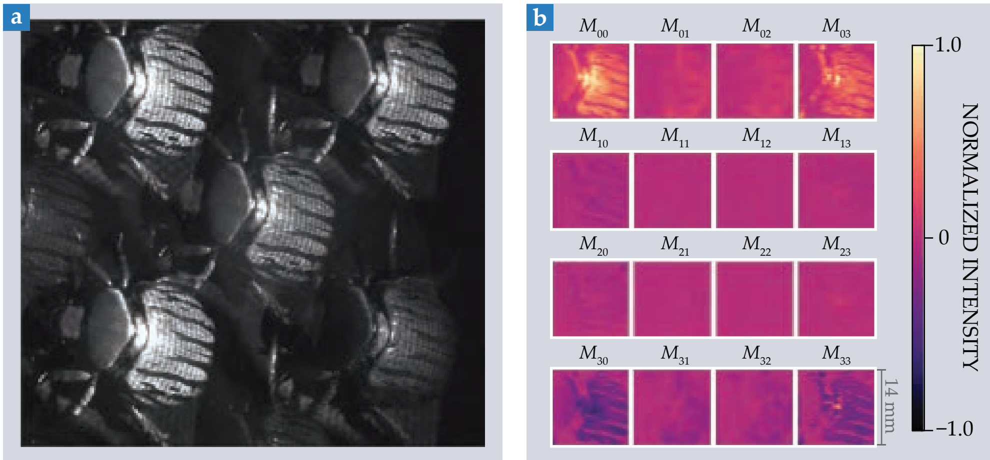


FIGURE 2. THE CHIRAL BEETLE, whose shell acts as a circular polarizer, is a convenient natural specimen for Mueller-matrix imaging. When the beetle was illuminated with structured light, the raw images **(a)** were obtained with the polarization analyzer in figure 1a. (The fifth image, in the center, is the zero-order, undiffracted light beam.) From those images, researchers calculated **(b)** the beetle's Mueller matrix. (Adapted from ref. 1.)

The idea was to bounce structured light off an object and then image it with a Stokes camera. Each pixel in the resulting Stokes images comes from light with a slightly different polarization; from the known structure of the incoming light, the researchers could calculate the object's Mueller matrix.

The structured-light approach has the slight disadvantage of sacrificing spatial resolution: If the object's Mueller matrix varies on a finer spatial scale than the structured light, those details are lost. But to the extent that the resolution loss is a problem, it can be solved by better engineering. "Fundamentally, our only limit is the diffraction limit," says Zaidi. "Everything else is fabrication: the size of the nanopillars, the size of the metasurface, and so on. There are lots of degrees of freedom when it comes to the design."

For one of their proof-of-concept experiments, the researchers imaged the chiral beetle—which, as the name suggests, has an unusual asymmetric effect on light. Figure 2a shows the four raw Stokes images. (The fifth image, in the center, is the zero-order, undiffracted light, which carries no additional polarization information.) Although too small to see, tiny fluctuations in the images encode the effects of all the polarizations of the structured light. From those, the researchers computed the Mueller matrix, shown in figure 2b.

The chiral beetle's shell is a uniform circular polarizer, so its Mueller matrix is nearly the same everywhere. The researchers expected—and found—the whole beetle to be visible in all the matrix's nonzero components. Furthermore, because it's a circular polarizer, they expected the only nonzero components to be in the corners: M_{00}, M_{03}, M_{30}, and M_{33}. The zeroth row and column carry information about the light's total intensity, and the third row and column encode its degree of circular polarization. The matrix elements in the middle, which encode the degree and direction of linear polarization, should be zero. That's more or less what the researchers found.

Wealth of data

"Mueller-matrix imaging existed already," says Zaidi, "but this work makes it more accessible. You no longer need bulky optical components at specific orientations that you need to control precisely. People in other fields, without experience in optics, might have shied away from that."

The potential applications abound in the biomedical and biochemical sciences. Nearly all biomolecules are mirror asymmetric, so they nearly all have some effect on circularly polarized light. Polarization measurements are a quick way of probing aspects of the composition—such as the sugar content—of a clear solution. It's even possible to noninvasively check diabetic patients' glucose levels by making polarization measurements on the clear parts of their eyes.[3]

And then there's cancer. Malignant tissues grow wildly and erratically, and their collagen fibers and other structures are jumbled up. So tumors have a different polarizing effect on light than do healthy tissues, whose collagen fibers are straight. Skin cancer is known to be detectible by Mueller-matrix imaging, but clinical implementation has struggled because of the complexity of the instrumentation.[4]

Skin cancer is already the easiest cancer to screen for by visual inspection. Most cancers grow in far less accessible areas, where neither doctors' eyes nor conventional optical instruments can go. But metasurface optics are small enough that they can reach parts of the body that are hard to access otherwise.

"We had a collaboration with a team at Massachusetts General Hospital to develop a metalens to put in an endoscope," says Capasso. "And for the first time, we could image the very beginning of a bronchial tumor, when its features were of wavelength size." That work didn't involve polarization measurements.[5] But the Mueller-matrix metasurfaces are no less compatible with endoscopic imaging.

Other applications may still be discovered. Mueller-matrix data are currently sparse, but as more images come in, researchers will learn more about how they can or can't use the data to classify materials, structures, and biological tissues. As Zaidi points out, the task is well suited to machine-learning algorithms. "Machine learning would find correlations that we're not aware of yet," he says. "It's easier for us to think physically in terms of polarizers and wave plates, but there may be other elements to the data that machine learning can pick up. I think there's a huge opportunity here."

Johanna Miller

References

1. A. Zaidi et al., *Nat. Photon.* (2024), doi:10.1038/s41566-024-01426-x.
2. N. A. Rubin et al., *Science* **365**, eaax1839 (2019).
3. J. S. Baba et al., *J. Biomed. Opt.* **7**, 321 (2002).
4. P. Ghassemi et al., *J. Biomed. Opt.* **17**, 076014 (2012).
5. H. Pahlevaninezhad et al., *Nat. Photon.* **12**, 540 (2018).

A new route to synthetic diamond

Most diamonds are formed at pressures exceeding 40 000 atm. With a new approach, nanocrystals can be grown in a liquid-metal mixture at just 1 atm.

Natural diamonds crystallize in Earth's mantle at high temperatures and pressures. The most popular method for making synthetic diamonds mimics those conditions by using anvil presses to achieve extreme pressures. Another approach, chemical vapor deposition (CVD), works at low pressures and is used to grow diamonds layer by layer on a substrate in a vacuum, a process more akin to the diamond growth that occurs in interstellar gas clouds.

Now Rodney Ruoff (at the Ulsan National Institute of Science and Technology and the Institute for Basic Science Center for Multidimensional Carbon Materials in South Korea) and colleagues have unveiled another route to diamond growth that works at ambient pressure. The approach, which involves methane and hydrogen gas and a catalyst of molten metals, requires less energy and less-advanced equipment than the two leading diamond synthesis methods.[1]

Below pressures of about 20 000 atm, graphite is the stable form of pure carbon. Diamond is kinetically stable at those lower pressures but not thermodynamically favored. CVD succeeds at growing diamond rather than graphite at low pressures by the inclusion of free hydrogen, which continually reacts with and etches away at any graphite that forms, while diamond is allowed to grow.[2] Many energy-intensive techniques—such as hot filaments, microwaves, or electric arcs—are used to ionize hydrogen gas and produce those free hydrogen atoms.

The new liquid-metal catalysis method differs from CVD in that it does not use any high-energy-density methods to ionize hydrogen gas, and the diamond is grown within a liquid-metal mixture, as shown in figure 1, rather than being deposited onto a solid surface from a vapor. And unlike many CVD approaches, the new technique does not rely on seed crystals. Ruoff says liquid-metal catalysis's lower energy requirement, seedless growth, and ability to spread liquid metal over large surface areas could make it an attractive alternative to the CVD and high-pressure, high-temperature (HPHT) synthesis methods, at least for some applications.

Seed of an idea

Ruoff was inspired to explore liquid-metal catalysis after learning of experiments in which scientists had grown single sheets of graphite from methane using liquid gallium as a catalyst.[3] Breaking the strong covalent bonds between the carbon and hydrogen atoms in methane usually requires a high activation energy, but gallium had significantly

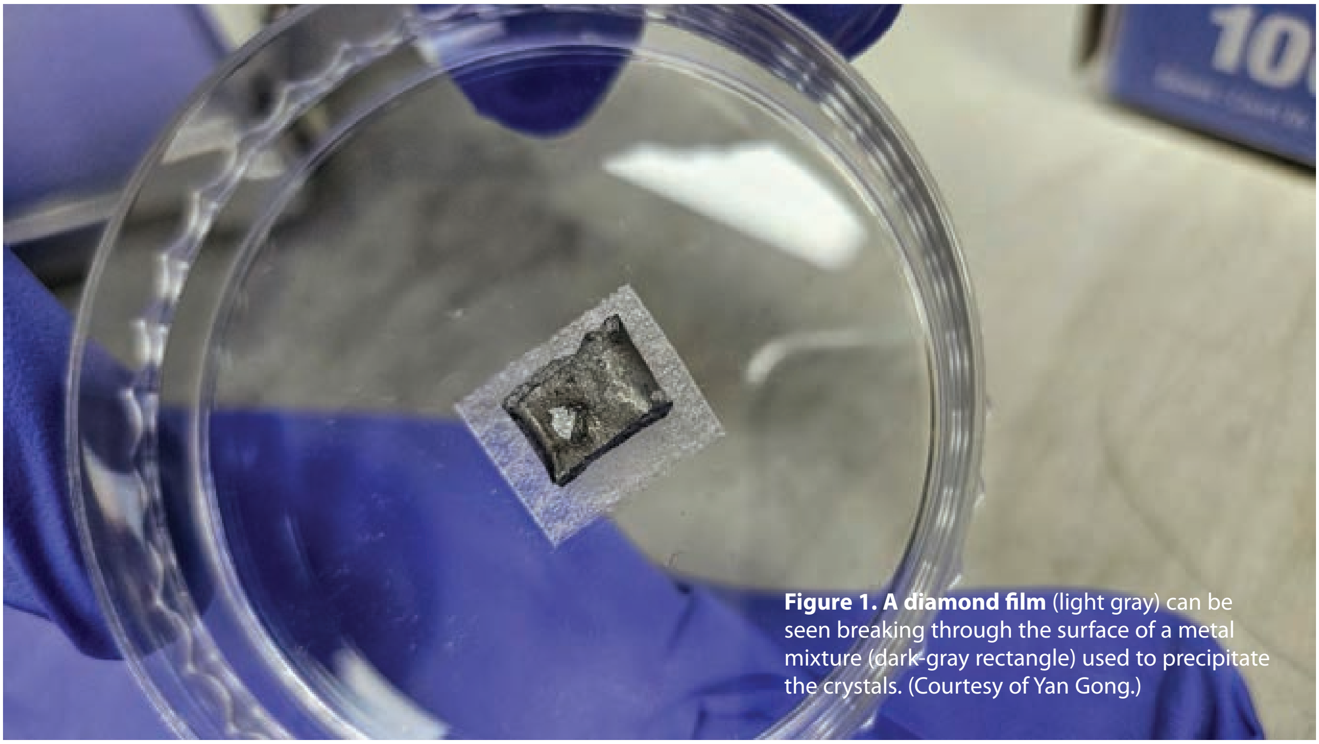


Figure 1. A diamond film (light gray) can be seen breaking through the surface of a metal mixture (dark-gray rectangle) used to precipitate the crystals. (Courtesy of Yan Gong.)

lowered the reaction barrier. And the catalysis had occurred at temperatures as low as 50 °C. (Pure gallium, with a melting point just under 30 °C, can turn to liquid from just the warmth of a hand.) The finding made Ruoff think that gallium's strength as a catalyst could lead to lower-energy diamond synthesis.

In early trials, Ruoff and his colleagues poured molten metals over diamond flecks in a chamber with carbon-bearing gases. They performed the experiments on silicon dioxide wafers, an unreactive base. The flecks didn't grow—until the metal spilled over the wafer during one experiment and touched pure silicon at its edge. Soon after, Yan Gong, a PhD student in Ruoff's lab, saw evidence of new pyramid-shaped crystal growth on the diamond seeds.

When the researchers looked closer, they found that new diamond had grown on the seed crystals and that the pyramids were topped by single crystals of silicon carbide. Their presence indicated that silicon from the wafer had dissolved into the metals and somehow aided the growth of diamond.

The researchers then pivoted to exploring the role of silicon, using a graphite crucible instead of a silicon wafer as a base. They experimented with varied concentrations of silicon in the mix and found that any concentration above 1% would grow silicon carbide. At smaller concentrations, they began to see diamond growth, but not in the way they expected. Graphite was growing on the seed crystals, but diamond was growing on the base of the graphite crucible. Again, the lab pivoted, this time to seedless growth.

In their new paper, Gong and colleagues now show how a mixture of gallium, iron, nickel, and silicon can form diamonds: When heated to 1025 °C in a chamber with methane and hydrogen gas, the mixture nucleated diamond nanocrystals at the base of the graphite crucible after 15 minutes and formed a continuous diamond film after 150 minutes (see figure 2). Although silicon's role in the nucleation process is not fully understood, it seemed to be crucial to the experiment's success. Silicon was incorporated throughout the lattice of the diamond, and slight variations of the silicon concentration produced different crystal sizes and densities. And when it was depleted from the mix, diamond growth came to a halt.

Carat and stick

The polycrystalline diamond film grown with the new method is far from the gem-quality single crystals that the jewelry industry prizes. As the method is further developed, it may get better at making larger individual crystals. But diamond's many valuable qualities—including extreme hardness, high thermal conductivity, and semiconducting electron transport—make it a useful material for a range of applications, not just jewelry. For example, diamond films are used as wear- and corrosion-resistant surfaces and as heat sinks on integrated circuits.

"Diamond is essential for everything in our daily life," says Soumen Mandal, a postdoctoral research associate at Cardiff University in the UK who specializes in nanocrystalline diamond growth. "For example, in your car there's almost a kilometer of wire, all drawn with diamond tools," he says. (For more on the emerging applications of diamond, see Physics Today, March 2022, page 22.)

CVD and HPHT are both commonly used to make gem-quality diamonds, for which color, clarity, and size are important. HPHT is used to make nearly 99% of synthetic diamonds and is the preferred method for industrial applications such as diamond drill bits and grit, for which the diamonds can be smaller and less pure than those used for jewelry. Diamond nanocrystals have also been made via ultrasonic cavitation on graphite and by explosions confined to metal cylinders to produce what's known as detonation nanodiamonds.

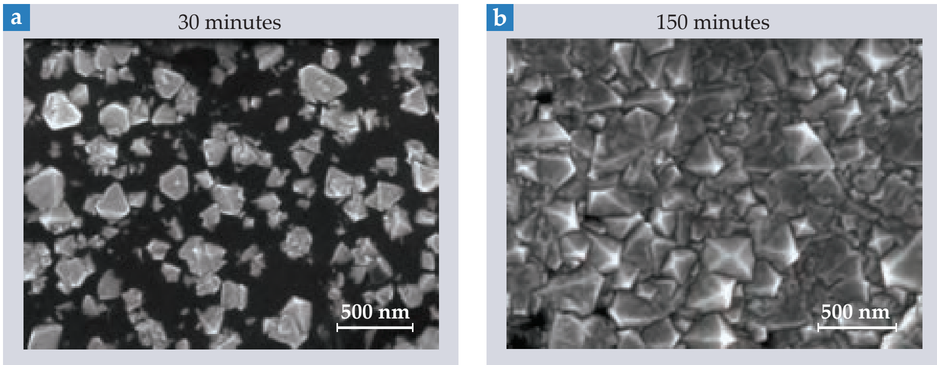


FIGURE 2. DIAMOND NANOCRYSTALS, captured in scanning electron microscope images after growing at ambient pressure for **(a)** 30 minutes and **(b)** 150 minutes, at which point they've formed a continuous film. Silicon aids the nucleation of the carbon crystals from a liquid-metal mixture. (Adapted from ref. 1.)

For research and technological applications, CVD is the preferred method, since its layer-by-layer growth allows for the greatest degree of control. To measure magnetic fields (see PHYSICS TODAY, May 2024, page 12) or build qubits, scientists can aim to create defects from specific atoms, such as nitrogen or silicon. In those highly sensitive applications, even the faintest bit of metal contamination will interfere with the intended use.

Mandal is somewhat doubtful that the liquid-metal catalysis approach will offer enough advantages to get widespread uptake. He says that both CVD and HPHT synthesis work well and have the benefit of several decades of research investment. Billions of carats of synthetic diamond are already produced every year, mostly via the HPHT method.

Ruoff plans to continue exploring liquid-metal catalysis for synthesizing not only diamonds but also materials made of other elements, such as boron and nitrogen. He and his colleagues are still unraveling the mechanisms that led to diamond growth in their experiment. Mixing a metal that doesn't dissolve carbon, such as gallium, with others that do, such as nickel and iron, may be an important factor. (Iron dissolves diamond so effectively that diamond is not used to cut it.) Hydrogen's role in the process also remains unclear.

Ruoff says that he hopes the result will inspire others to do basic research in the same vein. "I think that by exploring liquid metals broadly and with a lot of different configurations, new science will be done, even if people don't make diamond."

Laura Fattaruso

References

1. Y. Gong et al., *Nature* **629**, 348 (2024).
2. J. C. Angus, in *Diamond Films Handbook*, J. Asmussen, D. K. Reinhard, eds., CRC Press (2002), p. 17.
3. J. Fujita et al., *Sci. Rep.* **7**, 12371 (2017).

Ultralow-field MRI machine could cost less than a car

The prototype medical device, whose AI-processed images show features similar to those from a typical instrument with a strong magnetic field, could increase MRI access.

FIGURE 1. THIS MRI SCANNER operates at 0.05 T and costs much less than a conventional 3 T instrument. It can be placed in rooms that don't have expensive shielding, and its open design provides some level of comfort for patients. (Adapted from ref. 2.)

Many improvements in health care are the result of MRI. Doctors use the noninvasive imaging technique to diagnose strokes, brain injuries, tumors, and other conditions of the brain and spinal cord. The technique can image many other internal organs, and doctors often prefer it to other imaging techniques, such as CT scanning and positron emission tomography, because of MRI's lack of ionizing radiation.

Yet despite a half century of advances since the invention of MRI, the technology is still hard to access in low- and middle-income countries. The US has roughly 40 scanners per million people, but the average across all African countries is only about 0.7 scanners per million people.[1] A typical 1.5- or 3-tesla instrument uses a heavy superconducting magnet, which must be kept cool with expensive liquid helium, and power-hungry imaging electronics. In addition, the room housing the machine must be built with extensive shielding to block electromagnetic interference.

To lower the cost and power demand, MRI researchers have been developing alternative instruments that use simple permanent magnets and thus apply a relatively weak magnetic field. But the weaker fields lead to blurry images, which obscure the anatomical detail that doctors need when diagnosing and treating various diseases.

A research team led by Ed X. Wu of the University of Hong Kong has now constructed an MRI device that produces relatively high-quality images with a 0.05 T field, 1/60th that of a conventional instrument. The new machine uses the same amount of electricity as a hair dryer, and unlike previous low-field scanners, it's capable of imaging not just the brain but also other major organs. The researchers developed an artificial intelligence (AI) algorithm to sharpen the initially blurry images. Although the approach needs more clinical evaluation, the preliminary results show much of the same anatomical detail as images from a high-field instrument.[2]

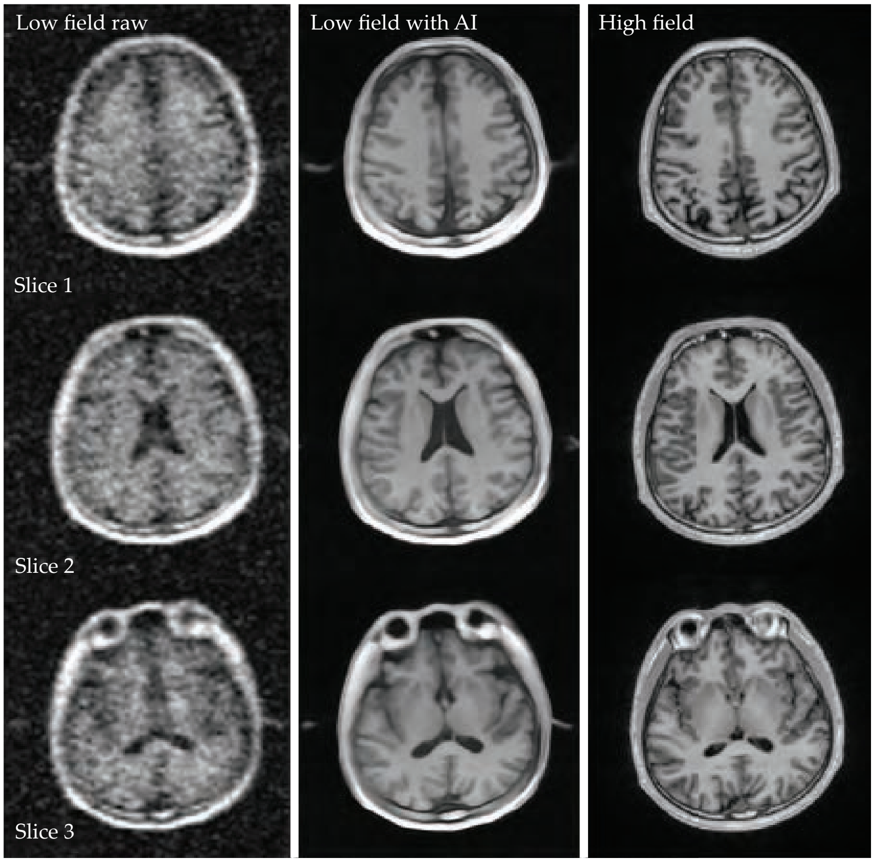


FIGURE 2. ARTIFICIAL INTELLIGENCE (AI) can reduce the noise and improve the spatial resolution of MRI. The left column shows image slices of a healthy human brain collected by a prototype 0.05 T scanner. The middle column shows the same images corrected with an AI deep-learning algorithm, and the right column shows results from a conventional 3 T scanner. Although some details are sharper in the high-field instrument's images, the AI-reconstructed images show many of the same anatomical features. (Adapted from ref. 2.)

Magnetic resonance

MRI scanners work according to the principle of nuclear magnetic resonance. Hydrogen nuclei in the human body, primarily in fat cells and water, become spin-polarized in the scanner's strong magnetic field, and in bulk the nuclei develop a small magnetization that's aligned with the field. Then a perturbing RF pulse pushes the bulk magnetization out of alignment so that it precesses about the applied magnetic field, similar to how Earth's rotational axis wobbles slowly over time.

The bulk magnetization doesn't stay out of alignment for long. Once the RF pulse is turned off, the magnetization returns to equilibrium in tenths to tens of seconds. As it relaxes, sensor coils detect the precessing magnetization. Applying small field gradients produces location-specific measurements that can then be used to construct an image. Nuclei in different tissues of the human body relax at different rates, so a particular tissue or organ can be targeted for imaging by applying certain sequences of specifically timed RF pulses. (For more on the physics of NMR and MRI, see Physics Today, December 2003, page 24, and the article by Dave Jordan, Physics Today, February 2020, page 34.)

MRI signal strength scales approximately with the square of the magnetic field strength, so low-field instruments start at a big disadvantage. Longer scans would increase the signal-to-noise ratio and thus could yield clearer images. But during an examination, which typically takes less than an hour, a patient must lie in a confined tube-shaped area and during each scan remain motionless—any movements introduce artifacts into the final images. The longer the scan, the more likely it is that a patient will move.

A low-field instrument won't make the examination time shorter, but it can be more comfortable for a patient. Because the permanent magnets don't need to be cooled with liquid helium, they can be designed with an open geometry, so the machine may feel less claustrophobic for patients. The scanner designed by Wu and colleagues is pictured in figure 1.

The more open design does have drawbacks. In addition to the RF noise that a patient might cause, low-field, unshielded MRI instruments are particularly susceptible to environmental sources of electronic noise such as computers and electric motors.

Extracting a signal

Wu and colleagues designed their MRI scanner without any RF shielding. Because it uses neodymium permanent magnets rather than a superconducting magnet, the device doesn't require much energy; the entire scanner can be powered by a standard electrical outlet.

A 2021 paper by the same group reported the key hardware components of a relatively small MRI device that could take images of the brain.[3]

To make the device capable of imaging not just the head, the researchers made a number of refinements to improve the magnetic field's homogeneity and limit unwanted eddy current in the magnets. The device—which is large enough to fit a patient's body between the top and bottom magnetic plates—is capable of imaging the head, heart, lungs, and abdominal area.

A lower-power machine, however, isn't worthwhile if it can't produce useful data. To get a clear image from an instrument with a weak signal and no shielding, the researchers first needed to understand how electromagnetic interference showed up in their signals. Using several coils placed around and inside the device, they measured both interference and the magnetic resonance signal during a normal scan and then measured only the interference when no magnetic resonance signals were collected. The data, in addition to high-quality images produced by high-field MRI instruments, were compiled to train an AI prediction algorithm that estimated interference-free MRI data.[4]

The algorithm uses deep learning, which is a type of AI that's loosely based on neural networks. It uses multiple classification layers to identify successively more complicated features in a data set. Because of advances in computing hardware and software, the method has become popular for all sorts of applications, including such complex tasks as predicting a protein's 3D structure (see Physics Today, October 2021, page 14) and forecasting the weather (see the article by Amy McGovern, Philippe Tissot, and Ann Bostrom, Physics Today, January 2024, page 26).

After training, Wu and colleagues' AI model reduces noise artifacts in the low-field instrument's images. Because the deep-learning algorithm is trained to identify the magnetic resonance signal, it can generate clean images even in a changing noise environment and without RF shielding. As shown in figure 2, the AI-processed images broadly capture the brain features that are visible in high-field images.

Wu and colleagues aren't the first to use AI to enhance the image quality of low-field MRI instruments. Several previous approaches focused on identifying and correcting for electromagnetic interference to improve brain MRI.[5] Unlike previous small devices, the one developed by Wu and colleagues is large enough to clearly image anatomical structures from multiple areas of the body.

Although the new results are improved, the low-field images and the high-field images do have differences—as seen, for example, in some of the smaller, darker features at the edge of the brain in figure 2. "This might be a problem for diagnostic radiology," says Mike Tyszka, the associate director of human neuroimaging at the Caltech Brain Imaging Center in Pasadena, "and would need careful evaluation in a clinical population."

Expanding MRI access

Wu and colleagues have begun some clinical evaluation. In two papers published last year, Wu's group reported that their low-field scanner revealed the same lesions in the brains of stroke patients as did a 3 T device.[4,6] Wu and others in the MRI community are working to determine the clinically relevant details that low-field MRI can identify and those it can't. "We need to improve the raw data quality as well as AI models," says Wu. "Both areas are still unsettled but full of hope, in my view."

Expanded MRI access is a big goal of the community, and low-field devices could help, if they're approved for broad imaging applications. Wu and colleagues estimate that if manufactured at an industrial scale, hardware material for their device would cost about $22 000. A reconditioned 1.5 T machine costs roughly 10 times as much, and a brand-new 3 T machine may be between $2 million and $3 million. The construction and installation expenses for an RF-shielded room can raise the price tag by a few million dollars more.

Both high-field and low-field machines need regular maintenance and repairs, so the total expense of the instrument developed by Wu and colleagues would be higher than the hardware cost. Still, the choice to use a permanent, low-field magnet and to forgo any RF shielding makes it far more affordable than a conventional, high-field MRI machine.

Because the low-field machine is small and doesn't need RF shielding, it could be useful in places that usually don't have an MRI machine, such as in an emergency room. Underserved communities in the rural US or other places worldwide would benefit from more MRI access. The most common type of liver cancer, for example, affects roughly 1 million people globally each year and is the fourth leading cause of cancer-related death worldwide.[7] If doctors could diagnose the disease at an earlier stage with the help of an affordable MRI scanner, a patient's prognosis could be improved.

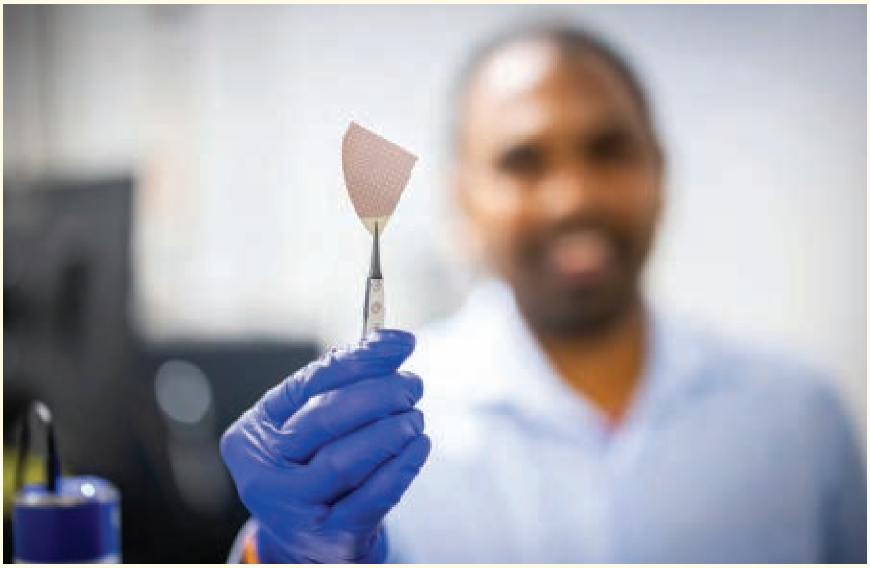
Alex Lopatka

References

1. U. C. Anazodo et al., *NMR Biomed.* **36**, e4846 (2023).
2. Y. Zhao et al., *Science* **384**, eadm7168 (2024).
3. Y. Liu et al., *Nat. Commun.* **12**, 7238 (2021).
4. C. Man et al., *Sci. Adv.* **9**, eadi9327 (2023).
5. S. A. Srinivas et al., *Magn. Reson. Med.* **87**, 614 (2022).
6. V. Lau et al., *Magn. Reson. Med.* **90**, 400 (2023).
7. J. M. Llovet et al., *Nat. Rev. Dis. Primers* **7**, 6 (2021).

New memory circuit can handle the heat

A ferroelectric device can store data at temperatures up to 600 °C for dozens of hours at a time.



DHIREN PRADHAN, a materials scientist at the University of Pennsylvania, displays a heat-resistant ferroelectric memory device. (Image from Dhiren Pradhan.)

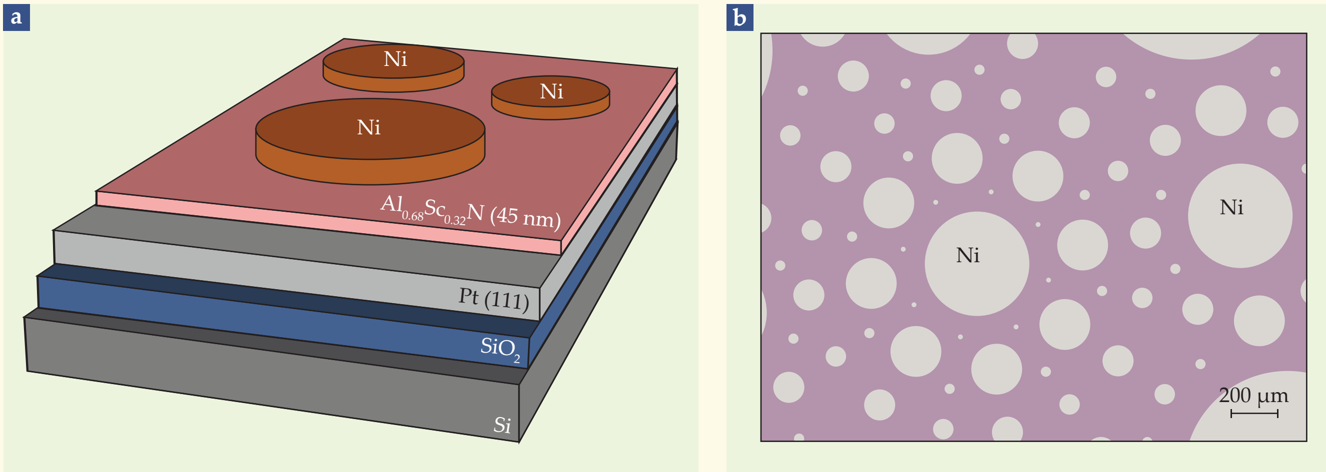
Silicon-based electronics stop working above 250–300 °C. That's the temperature at which the number of thermally induced charge carriers in a material exceeds its doping concentration. More generally, nonvolatile memory devices—circuits that can retain stored data even when unpowered—cannot reliably operate above that temperature. The limit puts devices made from CMOS materials out of range of the extreme temperatures found on the surface of Venus, for instance, or inside a jet engine.

Fortunately, a few wideband semiconductors, such as silicon carbide, can be the basis for logic transistors that operate effectively at up to 800 °C (see Physics Today, March 2017, page 19). But the absence of memory devices able to store information at elevated temperatures restricts the sophistication of any computing that can take place.

A group of electrical engineers and materials scientists led by Dhiren Pradhan (a postdoc at the University of Pennsylvania) and David Moore (at the Air Force Research Laboratory) has now demonstrated a scalable, high-temperature, nonvolatile memory device that can fill the gap. According to the group's measurements, the device, which consists of a 45-nm-thick aluminum scandium nitride ferroelectric diode, can reliably operate in temperatures up to 600 °C and quickly switch between stored states for more than 60 hours.

Shown schematically in the figure, the device's circular nickel electrodes are deposited atop the thin ferroelectric layer, which is grown on a 10 cm silicon wafer. Ferroelectric materials exhibit remanent polarization—meaning that when an electric field is applied and then removed, the material remains polarized—which can serve as binary states 0 and 1, the off and on states, respectively, of an electronic memory. The crystal structure of AlScN is advantageous for its strength and durability. The material's polarization is stable at temperatures as high as 1000 °C, and its surface can be made microscopically smooth and devoid of cracks, pores, and holes. Of the hundreds of devices that the researchers built and tested, only a small fraction were found to be faulty. (D. K. Pradhan et al., *Nat. Electron.* **7**, 348, 2024.)

R. Mark Wilson



A FERROELECTRIC MEMORY DEVICE, in schematic. **(a)** Aluminum, scandium, and nitrogen were sputtered onto a 150 nm film of platinum and silicon dioxide on a silicon substrate. Nickel-metal electrodes were then photolithographically deposited atop the AlScN. **(b)** A cross section of the device is shown as it would appear in optical microscopy. (Image adapted from D. K. Pradhan et al., *Nat. Electron.* **7**, 348, 2024.)

Old forests are irreplaceably cool

Satellite measurements confirm that the sudden disappearance of mature tropical forests has a more drastic effect on local land temperature than does the gradual growth of young forests.

SMALL CLOUDS pop up over Amazon forestland in this 2019 image from NASA's Aqua satellite. (Image by Lauren Dauphin/NASA Earth Observatory.)

A single wildfire or logging operation can destroy a centuries-old forest in a relative instant. Although recovering or newly planted forestland can counteract some of the deleterious climate effects of forest disappearance, established forests don't just spring up overnight. Now Yuxiang Zhang and Xuhui Wang (of Peking University in China) and their colleagues have quantified the implications of those contrasting time scales on the local surface temperatures of tropical forestland. The results imply a scenario that some climate models fail to capture: A fledgling young forest cannot provide sufficient local cooling to fully compensate for the temperature jump that follows the rapid disappearance of mature trees.

For their analysis, the researchers examined two decades' worth of land surface temperature data acquired by NASA's Aqua and Terra satellites. They overlaid those measurements on tree-cover maps, which they used to find tropical forestland and to identify areas where forests either emerged or disappeared.

When compared with nearby areas where forest coverage remained consistent, a region with forest loss had a surface temperature roughly 0.5 °C higher, the researchers report. Land with forest gain, however, was only about 0.1 °C cooler than a consistently forested region. Those results don't match up with those of multiple widely used climate-encompassing Earth system models, which tend to project deforestation and reforestation to have roughly equal and opposite effects on local temperatures. To investigate the potential mechanisms underlying their temperature findings, the researchers analyzed several satellite-measurable vegetation properties, including leaf area and canopy height, and found similar asymmetries between newly emerged and newly cleared forestland.

Because they made temperature comparisons among neighboring areas, the researchers may have missed larger-scale climate effects of forest gains and losses, such as changes in atmospheric circulation. Nonetheless, the results suggest that, at least from a local climate perspective, replacing a destroyed forest can negate only so much damage. (Y. Zhang et al., *Nat. Geosci.* **17**, 426, 2024.)

Andrew Grant

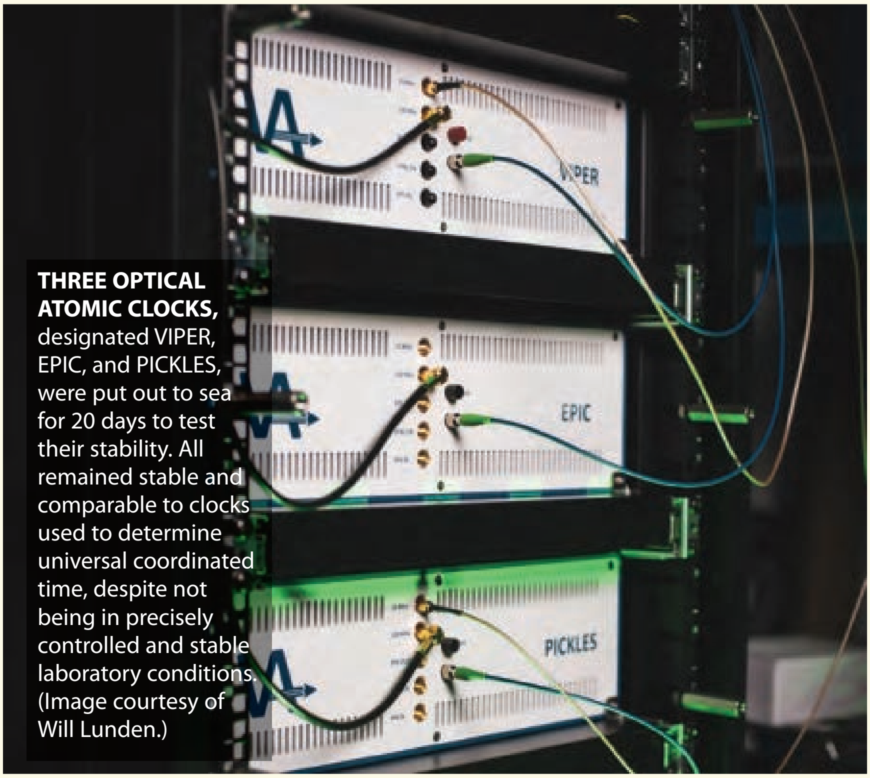
Keeping accurate time while on the ocean

Researchers use iodine to design smaller optical clocks for uses outside the laboratory.

Maritime navigation for centuries has been intrinsically linked with the accuracy of clocks. The 18th century saw many design competitions for seaworthy clocks that retained their accuracy enough to determine a ship's longitude. Nowadays, an accurate clock is important for participating in the Global Navigation Satellite System. Scientific and military vessels alike are equipped with microwave-based atomic clocks to achieve the precision necessary for safe sea travel. Such clocks, however, require a lot of external equipment to maintain their precision.

Optical atomic clocks, which provide higher performance than their microwave counterparts, have been primarily used in laboratory settings because they are bulky and sensitive to their surroundings. Researchers at Vector Atomic have now made a smaller optical clock that retains its precision outside the lab. The team, led by Martin Boyd, selected molecular iodine for the clock because its transitions are accessible with robust, industrial lasers. Several of its transition frequencies are also officially recognized as a length standard. Although using a molecule rather than an atom or an ion limits the accuracy of their clock, the researchers were aiming for a robust and mobile clock that is accurate enough for navigation.

The Vector Atomic clock differs from typical optical atomic clocks in that it uses a vapor cell, an iodine-filled tube with glass windows that allows for laser excitation and detection of the oscillations within. Because vapor-cell architecture is generally immune to vibration and orientation and, unlike many atomic clock designs, does not use consumables, it is an excellent choice for precision measurement in the field. It also doesn't require vacuum systems or cooling, thus further minimizing the accessory technologies needed to maintain the clock. Previous vapor-cell atomic clocks were only stable on a short time scale. The researchers counteracted that in part with a custom-built laser tailored for their clock.



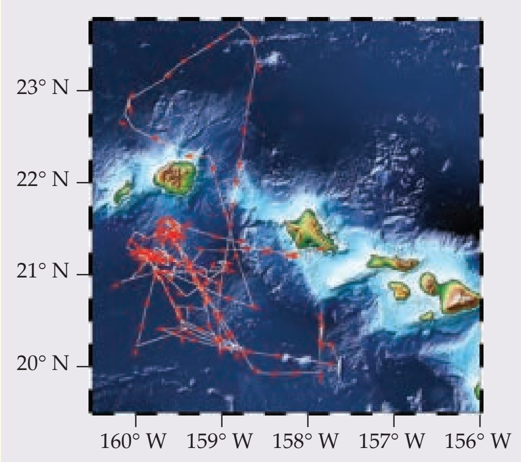
THREE OPTICAL ATOMIC CLOCKS, designated VIPER, EPIC, and PICKLES, were put out to sea for 20 days to test their stability. All remained stable and comparable to clocks used to determine universal coordinated time, despite not being in precisely controlled and stable laboratory conditions. (Image courtesy of Will Lunden.)

The team made three clocks to test. Two identical ones were tested against high-precision clocks at NIST and were found to be more stable than an active hydrogen maser over measurement intervals of under a thousand seconds. The frequency drift of the Vector Atomic clocks was only slightly greater over longer time scales than a calibration clock. The third clock had less stringent technical requirements but was designed for reduced size and power requirements. The researchers were satisfied with the lab tests, and the clocks were put out to sea.

For 20 days the three clocks were sailed around the Hawaiian Islands during a naval exercise. Despite unstable conditions at sea, the clocks remained accurate. Although not as precise as laboratory optical clocks, all three Vector Atomic clocks remained accurate enough to keep time within 400 trillionths of a second, even if access to an external clock was lost for a day.

The Vector Atomic team continues to make smaller, field-stable optical clocks. The researchers predict that the development of such clocks has the potential to reduce maritime reliance on GPS and also open opportunities for scientists to conduct high-precision studies at sea. (J. D. Roslund et al., *Nature* **628**, 736, 2024.)

Jennifer Sieben PT



A TEAM from Vector Atomic tested three optical atomic clocks on a naval ship that sailed around the Hawaiian Islands in 2022 as part of the Rim of the Pacific, the world's largest international maritime exercise. (Image adapted from J. D. Roslund et al., *Nature* **628**, 736, 2024/Vector Atomic.)

What is nuclear energy's role in mitigating climate change?

Cost, construction time, and safety, security, and proliferation risks all figure in.

The US and two dozen other countries have pledged to triple the world's nuclear energy capacity by 2050. Launched last fall at the United Nations Conference of the Parties (COP 28) in Dubai, the pledge is intended to help reach the goal of net-zero greenhouse gas emissions and limit global warming to 1.5 °C above preindustrial levels.

But is such a large increase in nuclear energy production feasible? Skeptics say that building nuclear reactors is too slow and costly to effectively mitigate climate change. And they say that security, safety, and proliferation risks need to be assessed in the context of today's geopolitics. Proponents say that nuclear energy is necessary in the climate change equation and that to wield influence in the nuclear arena, the US and other Western nations must be at the forefront of nuclear energy development and exports.

Kathryn Huff was assistant secretary in the US Department of Energy's Office of Nuclear Energy until May, when she rejoined the department of nuclear, plasma, and radiological engineering at the University of Illinois Urbana-Champaign. "We cannot meet our net-zero goal for the whole economy by 2050 without significant increase in nuclear power," says Huff. "It's not a statement of what is likely or probable. It's a statement of what is necessary."

Time, money, and nuclear energy

About 440 nuclear power reactors operate in 32 countries and Taiwan. They provide roughly 9% of electricity globally; in the US, that number is around 19%. China is building reactors at the fastest rate. Russia is the largest exporter of nuclear reactors; it is selling and setting them up in Egypt, Turkey, and other countries. Two commercial nuclear power reactors went on line in the past year at Plant Vogtle in Waynesboro, Georgia, bringing the US total number of operating reactors to 94.

In the drive to triple nuclear energy, some governments are giving much attention to small modular reactors (SMRs), which would produce a few hundred megawatts, making them about one-third the power of conventional gigawatt-scale reactors. Their appeal lies in the assumptions that they could be manufactured in assembly-line mode, which would keep costs down; could be distributed widely even to small users; and would have limited radiological release in an accident because of their size. Utilities or other customers could add to their stock of reactor modules as needed.

Historically, reactor projects in the US and other Western countries have been plagued by delays and cost overruns. The Vogtle reactors, for example, started up seven and eight years late and more than doubled in cost, from an initial estimate of $14 billion to a final cost of $34 billion. Ongoing projects in the UK, Finland, and France—the poster child for nuclear energy—are similarly late and more expensive than planned.

Ted Jones is senior director for national security and international programs at the Nuclear Energy Institute, a US-based nuclear industry trade association. The lowest-cost route to reducing greenhouse gas emissions involves nuclear energy, he says. To reduce US emissions by 95% by 2050, nuclear energy should be increased to provide 43% of US electricity needs, according to models he cites by the company Vibrant Clean Energy. The models also expand the contributions of wind and solar energy and battery storage. Tripling nuclear electricity production requires rebuilding the supply chain and stopping the cost and time overruns associated with reactor construction. "It will be hard," he says, "but it's realistic to believe it will improve."

More reactors, more targets

Sharon Squassoni is a former US State Department analyst who is now a research professor of international affairs at George Washington University. The


INTERNATIONAL ATOMIC ENERGY AGENCY/CC BY 2.0

THE ZAPORIZHZHYA POWER PLANT in Ukraine has been targeted by Russia during the war. The incidents at the facility highlight the specter of increased potential for attacks—military and terrorist—on nuclear plants if more are built to tackle climate change.

pledge to triple nuclear energy, combined with Russia's attacks on the Zaporizhzhya nuclear power plant in Ukraine, prompted her to write the report *New Nuclear Energy: Assessing the National Security Risks*, which came out in April. More reactors around the world, she says, means more potential targets. And the danger is enhanced if those targets are in countries that have unstable governments.

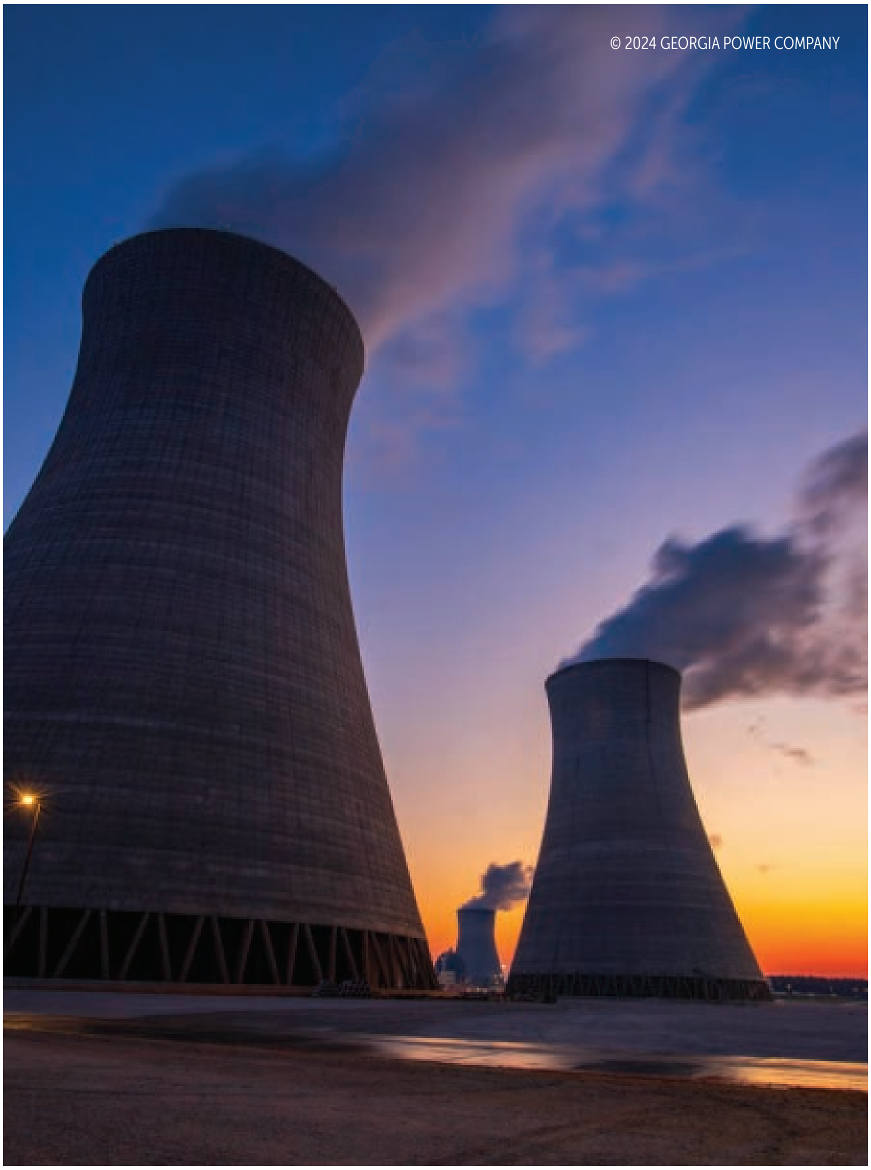
In her report, Squassoni urges the US government to convene an international study on the national security risk of SMRs. She also says that the State Department should commission its International Security Advisory Board to study how national security risks posed by nuclear energy have changed over the last two decades. In addition to proliferation risks, she says, the study should assess nuclear terrorism, sabotage, and weaponization of nuclear power plants. She also recommends that the US weigh nuclear solutions to climate change against other low-carbon options.

Countries new to nuclear reactors will need to train workers. And the know-how and the access to uranium fuel could be diverted to weapons purposes, says Henry Sokolski, who previously worked at the Pentagon and is now executive director of the Nonproliferation Policy Education Center. He calls nuclear power plants "bomb starter kits."

Economics and geopolitics

Dozens of SMR designs exist. They use various coolant types, including light water, liquid metal, high-temperature gas, and molten salt. For now, says Mark Jacobson, a professor of civil and environmental engineering at Stanford University, SMRs are still "vaporware. They don't exist." He and others note that historically, the size of reactors increased to get more electricity per dollar invested. Claims that the cost of electricity per plant will go down with SMRs "have not been validated," says Sokolski. Last year NuScale Power's plans to build a set of SMRs in Idaho to serve municipal utilities in Utah fell apart after the projected cost tripled.

Economics is what makes reactors so hard to realize, says Peter Bradford, who served on the Nuclear Regulatory Commission from 1977 to 1982, has chaired state utility regulatory commissions, and has taught courses on nuclear law and energy policy. The industry and the US government have a pattern, he says: "Every time a promised nuclear renaissance fails, they come up with some other reactor concept. SMRs are just the latest. But they never solve the cost problem." Still, governments and the nuclear industry remain eager to commit immense sums of taxpayer and customer money, he says. "I scratch my head at that."



COOLING TOWERS at the Vogtle plant in Georgia, where two new reactors went on line in the past year. The plant hosts 4 of the now 94 operating reactors in the US. The reactors' huge time and cost overruns exemplify challenges facing the expansion of nuclear capacity.

Many physicists support nuclear energy, says M. V. Ramana, a professor at the University of British Columbia's School of Public Policy and Global Affairs. His focus is on nuclear energy, especially SMRs, and he has written a forthcoming book on nuclear energy and climate change. He was the lone critic on a panel discussion about SMRs at the American Physical Society's April meeting, he says, and many in the audience were "less than open" to his views. He surmises that physicists "have a fundamental belief that the technology used to make nuclear weapons must also have a good use and that 'we have to redeem ourselves by taming the atom.'"

Given the costs of reactors and the snail's pace of construction, the tripling of nuclear energy is not going to happen, says Ramana. "It's moot." Instead, he sees the focus on nuclear energy as a distraction. "From the viewpoint of climate change," he says, "reactors are a diversion, and the money from the government is money that could go to renewables and to energy storage." At COP 28 in Dubai, 133 countries, including the US, committed to tripling the world's installed renewable-energy generation by 2030.

But US commitments to build nuclear reactors are motivated both by climate change mitigation aims and by geopolitical influence. At a 23 April press conference on Squassoni's recent report, Jane Nakano, a senior fellow in the program for energy security and climate change at the Center for Strategic and International Studies, said that for national security reasons and political influence, the US may have no choice but to pursue SMRs.

"If the US fails to build reactors, we will not only fail to meet climate goals, but we may cede our nuclear energy leadership to our adversaries," says the University of Illinois's Huff. "That does have real risks." Leadership in nuclear technology allows the US to drive the global conversation about safety, safeguards, and security, she explains.

Ramana disagrees: "Such zero-sum thinking will ensure that the climate crisis becomes worse."

Toni Feder

Einstein statue unveiled in Havana

Albert Einstein visited Cuba briefly in 1930. This past March, he came back to stay—in bronze.

Ernesto Altshuler, a physics professor at the University of Havana in Cuba, got the idea for a statue of Albert Einstein while he was planning celebrations for the International Year of Physics in 2005. It was the centenary of Einstein's annus mirabilis, when he published four groundbreaking papers. "I thought that a statue would be interesting both intrinsically and to attract students," Altshuler says.

It took nearly two decades, but a life-size statue of Einstein now graces the entrance to the university's physics department. It was inaugurated on 27 March.

The statue took so long to realize, Altshuler says, mostly because he wasn't able to raise money for it. In late 2005 Altshuler took to the Web to vent his frustrations with a post he titled "A dream that didn't come true."

A dozen years later, in 2017, that vent caught the eye of Wolfgang Bietenholz, a physicist at the National Autonomous University of Mexico. "I thought the statue was a nice idea, and it was a pity if they couldn't build it for lack of money," he says. Bietenholz is originally from Switzerland and, as he puts it, "knew about foundations there that were looking for reasonable projects." He played matchmaker, and an undisclosed foundation paid for the statue.

Once the money was secured, delays continued because of bureaucracy, the COVID-19 pandemic, and a dearth of materials needed to make the statue.

Havana sculptor José Villa Soberón took on the project. He is known internationally for statues of John Lennon, Gabriel García Márquez, Napoleon Bonaparte, and many others. Altshuler stood in as a model for Einstein's body. The hat Einstein holds is based on a gift he received during his 1930 visit.

In a sweet twist, Diego Valdés, a physics student from the University of Havana, helped cast the statue. Villa Soberón had turned for help to Valdés's father, a local sculptor who has a foundry. Valdés was involved throughout: casting wax, coating the wax cast with silica sand and plaster, melting the wax, pouring melted bronze into the resulting hollow mold, and touching up the details. The sculpture was cast in six pieces and later welded together, he says. "I thought it was great that we could cast a sculpture destined for my faculty at the university."

The statue is attracting attention, says Altshuler. "Einstein is one of the few persons in modern history that everyone can recognize."

Toni Feder

ALBERT EINSTEIN is the newest fixture at the physics department at the University of Havana in Cuba. Physics student Diego Valdés (next to Einstein) helped cast the statue, which was the brainchild of physics professor Ernesto Altshuler (left). (Courtesy of Ernesto Altshuler.)

Arecibo STEM educational center to open soon

Biology and computer science activities replace the iconic radio telescope at the Puerto Rican observatory site.

AMANDA JOY MEYERS/NSF

THE ÁNGEL RAMOS FOUNDATION Science and Visitor Center at the former Arecibo Observatory site is being converted into an educational center, scheduled to open in November.

Two years after the 305-meter radio telescope collapsed at Arecibo Observatory on 1 December 2020, NSF announced that the site would be converted into an educational center (see "NSF's Arecibo strategy puts future research into question," Physics Today online, 15 November 2022). NSF's Arecibo Center for Culturally Relevant and Inclusive Science Education, Computational Skills, and Community Engagement (NSF Arecibo C3) is scheduled to open in November. The move is simultaneously welcomed by and disappointing to people who visited or used the observatory for research.

NSF Arecibo C3 will feature a laboratory and an interactive science center with STEM exhibits and activities focused on biology and computer science. The center will target educators, K–12 students, and the general public.

A pilot phase began in late May. High school and college faculty from Puerto Rico and the rest of the US were invited to a workshop for a deep dive into nanopore DNA sequencing, a process used to determine the sequence of nucleotides in a DNA strand.

At the time Physics Today went to press, the schedule also included two biology-education workshops in which high schoolers and middle schoolers could extract, analyze, and model DNA. The goal of those activities is to expand students' understanding of genetics and cell biology.

Professional development for life sciences instructors will be a focus moving forward, says NSF Arecibo C3 lead investigator Jason Williams. A group of Puerto Rican high school teachers is helping to shape the center's programming. The center's executive director, Wanda Liz Díaz Merced, worked with the Puerto Rico Department of Education to incorporate astronomy learning for teachers and students.

Another pilot-phase activity hosted by the center is a citizen science project on biodiversity in which volunteers across the island collect samples of various plants, insects, and small animals and map their distribution.

Over the past two decades, NSF has invested more than $200 million in Arecibo operations, management, and maintenance. In 2023 NSF said that it is putting $5 million over five years into the center. Four institutions are partnering with NSF to organize and implement educational programs with

THE COQUI FROG is among Puerto Rico's indigenous plants and animals studied in the biodiversity mapping project run by the new educational center at the site of the former Arecibo Observatory. (Photo by Louis Porras; from V. H. González-Sánchez et al., *ZooKeys* **1022**, 79, 2021.)

that funding. The University of the Sacred Heart in San Juan will incorporate astronomy topics and create STEM activities for those with disabilities. The University of Maryland, Baltimore County, will focus on computational science. New York's Cold Spring Harbor Laboratory will promote biology through hands-on activities. The University of Puerto Rico at Río Piedras will offer undergraduate biology education through research and connect the University of Puerto Rico system to the center.

During its long run as the world's largest radio telescope, the Arecibo Observatory was used to detect hundreds of near-Earth asteroids, discover the first binary pulsar, and find the first exoplanet. The dish faced years of deterioration but continued to be used for many other significant discoveries before it collapsed (see "NSF puts Arecibo Observatory on chopping block," Physics Today online, 24 November 2020). NSF will highlight the observatory's legacy in an exhibit at the center.

A 12-meter radio telescope at the Arecibo site stopped operations last year when the facility officially closed. Astronomer Abel Méndez at the University of Puerto Rico at Arecibo's Planetary Habitability Laboratory trained students on the dish. He says that he hopes that funding will be found to bring the telescope back on line for educational and research purposes at the center. So far, though, that's not in the plans, according to an NSF spokesperson. Although gathering new astronomical data is not on the horizon for Arecibo, archived data from the telescope may be used in center activities.

Anish Roshi, who headed radio astronomy at Arecibo, says that a center for science education will benefit Puerto Rico. Still, he says, without a major research facility, visitors won't be exposed to world-class science.

Hannah H. Means

Indigenous women thrive in Los Alamos internships

A small program is having a big impact on its participants.

When Victoria Nofchissey spotted a flyer for a program called Engaging Indigenous Women in Nuclear Physics, she thought, "That's me!" She was a sophomore at Fort Lewis College in Durango, Colorado. Growing up without water or electricity on the Navajo Reservation in Arizona, she says, "I didn't imagine I had the possibility of being a physicist."

ELAINA SALTCLAH

ELAINA SALTCLAH (left) and Arielle Platero (right) went to Los Alamos National Laboratory as part of the Engaging Indigenous Women in Nuclear Physics pilot program. Bade Sayki (center) is a student at the University of Colorado, Boulder, who is working full time at the lab as a graduate research assistant.

A few months later, in the summer of 2023, she was at Los Alamos National Laboratory.

Indigenous women in STEM "are the minority of minorities," says Cesar Da Silva. (According to data from the Department of Education, Indigenous women earned 0.2% of bachelor's degrees in the physical sciences in 2021. See the interactive tool at https://ww2.aip.org/statistics/physics-engineering-degrees-earned.) He and fellow Los Alamos physicist Astrid Morreale proposed the program in 2021 in response to a call by the Department of Energy to increase participation by groups historically underrepresented in science. Their proposal to mentor such students won $200 000 for a pilot program, plus $484 000 from other sources. The program is now up for renewal funding.

The program has grown each year, from two to four to—this year—six participants.

Students spend 10 weeks in the summer at Los Alamos. They code, analyze data, and build detectors. They attend and present at conferences. Some visit other national labs. They meet scientists at Los Alamos and see first-hand how science is done.

Nofchissey's internship took her to CERN for three weeks. "I was most interested in glueballs," she says. "We look at collisions between protons and lead, lead and lead, and protons and protons to figure out if glueballs exist." Her project is part of the CERN LHCb experiment, which studies the bottom quark to gain insights into the asymmetry of matter and antimatter in the universe.

Being part of a 1500-strong international collaboration has been "eye-opening," Nofchissey says. "Once you specialize, everyone knows each other. It makes me happy to know there is a community out there."

Elaina Saltclah is also Navajo and a physics major from Fort Lewis College. As an intern at Los Alamos in summer 2023, she wrote code to analyze proton–proton collisions using spherical harmonics. The internship, she says, made it clear that "this is what I want to do with my life."

"My first encounters at Los Alamos went over my head," she says. "But people took time to take us aside and teach us the vocabulary. The environment fosters learning."

Importantly, Saltclah says, "the program allows us to be who we are. That means it can foster people who would otherwise be overlooked."

Being aware and accepting of the students' cultures is key, says Da Silva. He notes that the Indigenous students tend to be tightly connected to their families and communities. Some of them have kids. Few have been out of the area, let alone out of the country, he says. And, for example, none of the Indigenous women worked the week of the 8 April solar eclipse "for cultural reasons."

Da Silva's goal is to keep mentoring students, even beyond the summer program. "I want to make sure all of them continue in their careers," he says. "It doesn't have to be in physics. I try to help them apply for jobs in labs and for PhD studies. That is part of the program too."

DOE and DEI

The Los Alamos program is small, but it's a great example of what DOE is trying to do, says Harriet Kung, acting director of the agency's Office of Science. "Physics can only excel by including more students with diverse backgrounds. They bring diverse experiences and new ways of approaching scientific problems."

Other DOE programs in the diversity, equity, and inclusivity (DEI) space include a two-week virtual wildlife simulation camp aimed at students from Indigenous backgrounds and a partnership between Los Alamos and Navajo Technical University in Crownpoint, New Mexico, to train students to detect contamination from uranium mines and to build a particle detector for the LHCb.

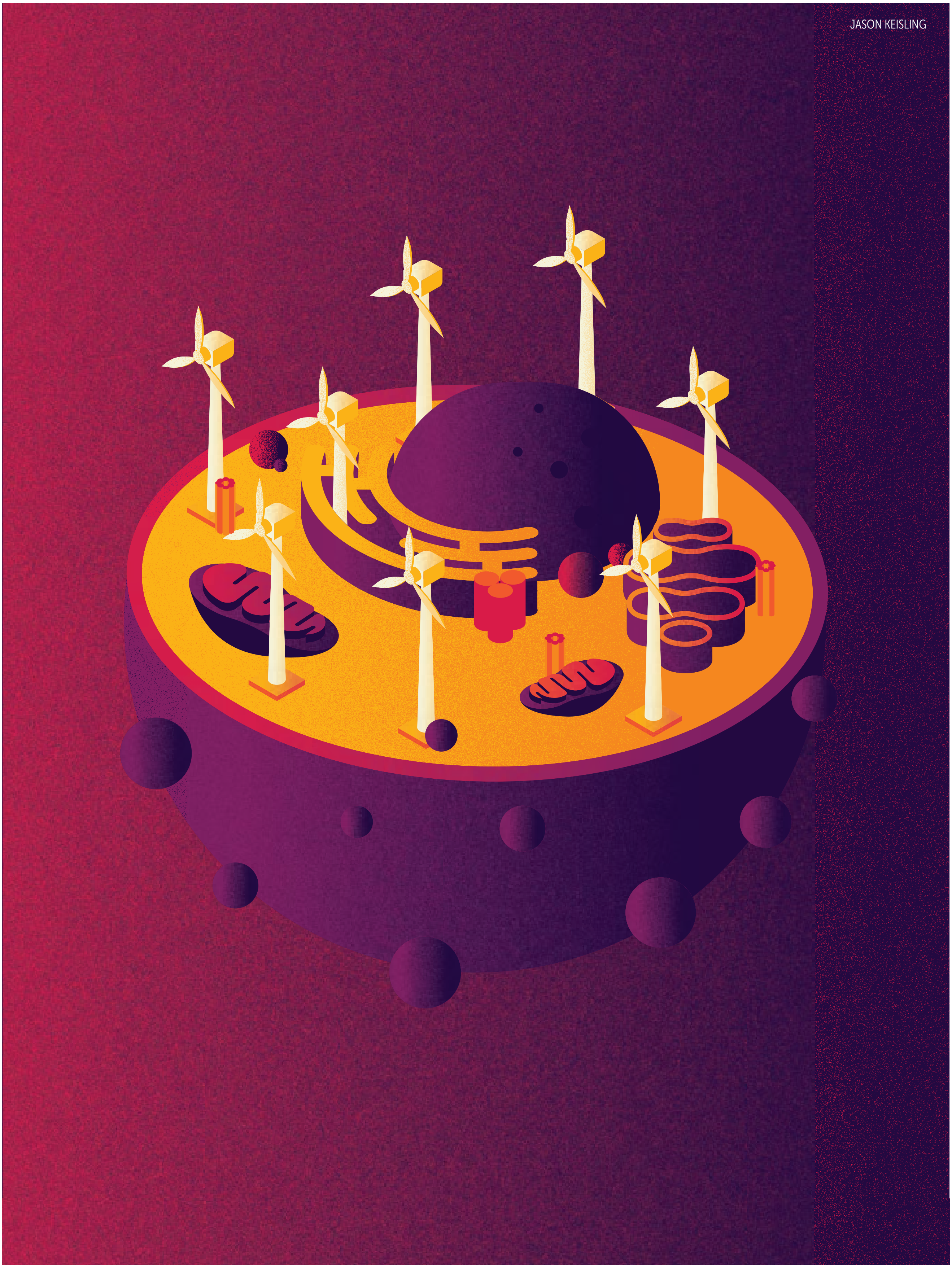
On much larger scales, in the past couple of years, DOE launched the Reaching a New Energy Sciences Workforce and Funding for Accelerated Inclusive Research programs to support individuals and institutions that are historically underrepresented in STEM. (See Physics Today, May 2024, page 28.) In total, the agency invests nearly $84 million annually in initiatives that have a DEI emphasis.

The federal government's efforts contrast with the anti-DEI legislation that is popping up in states around the country (see Physics Today, April 2024, page 22). How the state laws may affect programs like the Los Alamos internship for Indigenous women is unclear, but DOE is keeping an eye on whether such legislation could hinder advertisement of or applications to the program. To date, neither Colorado nor New Mexico has introduced anti-DEI legislation.

Until recently, Fort Lewis College was the sole source of physics students for the Los Alamos program. The college is a federally designated Native American–serving nontribal institution and the highest producer of Indigenous STEM bachelors in the US. This year, two students from Navajo Technical University will join the program.

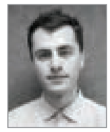
"For women in STEM, especially minority students, being able to see themselves in a career path is important," says Laurie Williams, an engineering professor at Fort Lewis College. "I have seen such a difference in the students after just one summer in the Los Alamos internship program. They have more confidence and higher aspirations." The internship is hitting all the high points for transforming their trajectories, says Williams. "The students feel they are having an impact. That is empowering."

Toni Feder PT



JASON KEISLING

Annette Kim is an undergraduate student concentrating in human developmental and regenerative biology with a secondary focus in neuroscience at Harvard College in Cambridge, Massachusetts. **Tyler Alexander**, interested in problems at the interface of medicine and biophysics, is an MD–PhD student at Thomas Jefferson University in Philadelphia. **Harvey Rubin** is a professor of medicine with a secondary appointment in computer and information science in the school of engineering at the University of Pennsylvania. **Divyansh Agarwal** is a clinical fellow in surgery at Harvard Medical School and an affiliate member of the Massachusetts General Hospital Center for Environment and Health.

WHEN CELLULAR SYSTEMS MEET POWER GRIDS

Annette Kim, Tyler Alexander, Harvey Rubin, and Divyansh Agarwal

Although biological energy systems and electrical grids differ in scale and are studied by different disciplines, the strategies from one system could lead to benefits for the other.

Most of us don't think twice about where our power comes from when we flip a switch to turn on the lights or plug in our electronic devices to charge. With our ever-increasing reliance on electricity to power our modern comforts, the task of improving the efficiency of electricity generation, the responsiveness of power grids to changing demands, and the system resiliency in the face of unexpected disruptions is crucial. Furthermore, outages in the traditional US power grid—made up of the Eastern, Western, and Texas power grids—are estimated to cost businesses upward of $150 billion a year.[1] Recent global trends toward upgrading to a predictive smart grid system and the rising popularity of renewable energy hold immense space for innovation.

Luckily, in addition to already established mechanisms of energy generation, another source of inspiration can potentially help engineer improved electrical grids: nature. Living systems have developed over eons of evolutionary pressures to adapt to fluctuating energy sources, transform multiple types of energy for both immediate and long-term uses, distribute energy in the form of specific compounds throughout intracellular compartments, and use that energy to signal their internal energy state to other members of their environment.

Energy in cells is generated through a process called glycolysis, in which carrier molecules are transmitted to the mitochondria—the powerhouse of cells. Both oxygen and energy are created, the latter in the form of adenosine triphosphate (ATP), which can be put to use wherever it is needed. If there is an excess of energy, it can be stored as glycogen.

Surprisingly, modern electrical grids closely parallel those processes, which can be most clearly seen when considering the five key elements of engineered energy grids: the generation

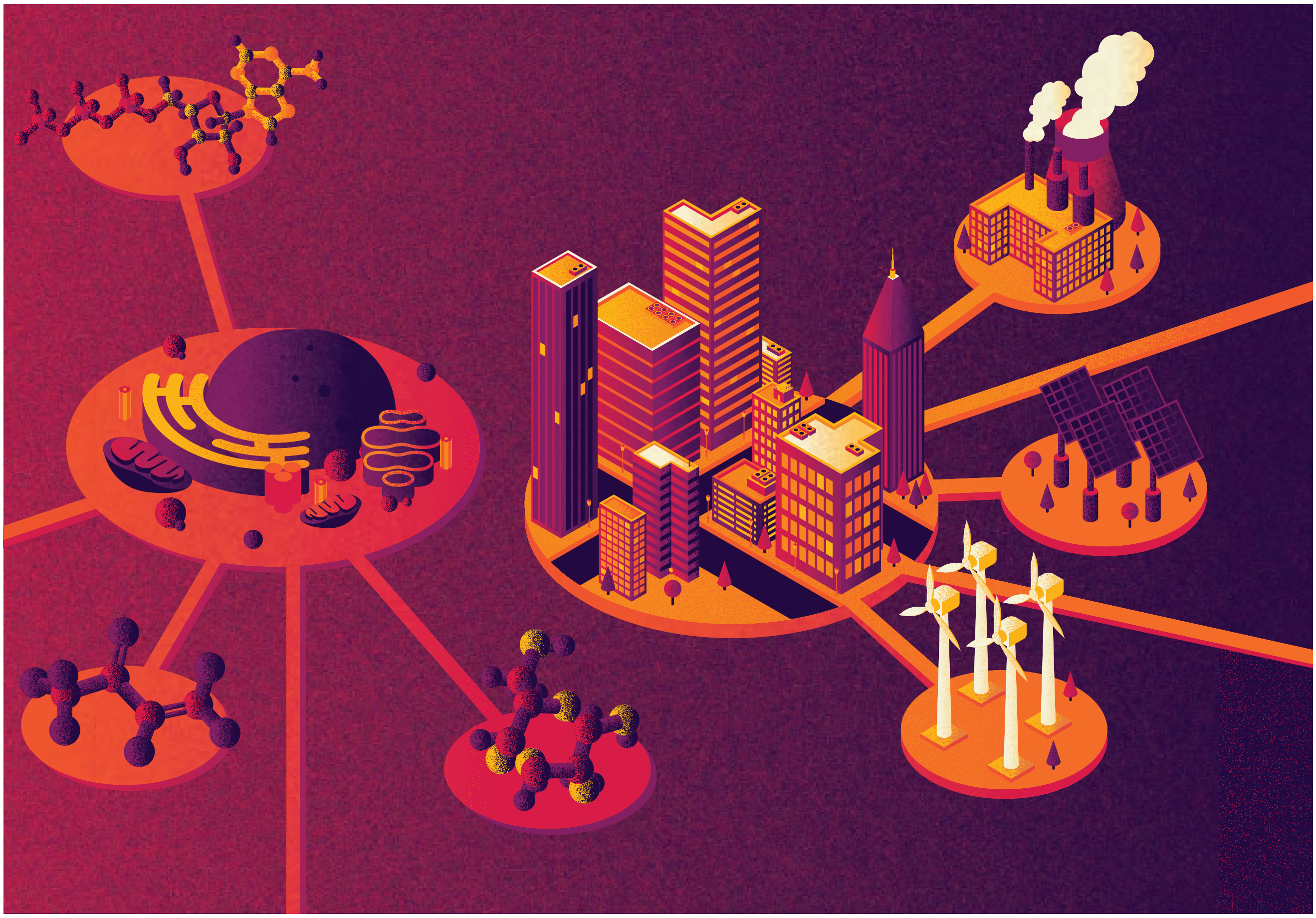


FIGURE 1. DIFFERENT ENERGY SYSTEMS, such as a metabolic pathway in the body (left) and a modern electrical grid (right) are more similar than they might first appear. Both use multiple sources of energy generation and dynamically respond to changing demands by altering energy pathways. Improvements in electrical grids could be inspired by looking at how cells manage a body's energy. (Illustration by Jason Keisling.)

source, the transmission system, the substation for sharing and redistributing the energy, the storage reservoir, and, ultimately, the consumer.

Investigating those parallels could lead to improved energy grids. Biological systems demonstrate a way to optimize storage, increase resilience, respond to stochasticity, and improve coordination between multiple energy types that isn't yet seen in electrical systems. That functionality is especially important in the face of accelerated climate change and an increased dependence on renewable energy—whose availability may not always be reliable. Optimal control algorithms for stochastically fluctuating data are needed to avoid power imbalances, failures that immediately cause failures of their dependents, and unstable energy transmission. The future of the energy field is in renewable energy and more resilient grid systems. To achieve that future, we explore the ways electrical engineers can learn from biological systems.

Energy through multiple pathways

As renewable-energy sources, such as solar and wind, become increasingly popular, the ability to harness electricity from different resources can improve the reliability of the US power grid.[2] Living systems already intrinsically synthesize energy from multiple sources. For instance, at a macroscopic level, honey, a slice of bread, and a banana provide different forms of energy, but they can all be converted into ATP through cellular respiration. The ways in which energy sources are used to produce electricity are also distinct. Solar power, for example, directly excites electrons in a photovoltaic cell, whereas wind energy turns a turbine to generate electricity. Other sources of direct energy, such as nuclear power, coal, and fossil fuels, share a similar energy-generation pathway to ultimately produce electricity. A living organism, however, operates similarly to newer smart grids in that not only are the sources of energy distinct, as shown in figure 1, but the processes by which energy gets converted to a usable form are also distinct.

Biology employs a series of connected chemical reactions, called metabolic pathways, to convert a molecule into an end product. Different forms of carbohydrates, for instance, move through distinct metabolic pathways in the presence and absence of oxygen and varying substrates. Proteins are broken into their constituent amino acids and enter the metabolic process at different points, depending on the amino acid. And lipids are metabolized through yet another pathway. Ultimately, those pathways converge into a single form of energy: ATP. Yet the pathway through which a form of energy travels in a cell is largely determined by metabolic demands and regulators. The intrinsic ability of biological systems to rapidly switch their source and mechanism of energy production, depending on the host and microenvironment factors, can provide insights relevant for augmenting the self-sufficiency and resilience of existing engineered grids.

Both systems work best when they have a greater likelihood of being able to provide energy when one or more sources become unavailable. Whereas biology has the advantage of billions of years of evolution to integrate its diverse metabolic pathways, the integration of more renewable-energy sources into our existing power grids is an ongoing challenge.

Energy consumption during summer months in the US peaks in the afternoon when temperatures are hottest. But during the winter months, there are two peaks of lesser magnitude in the morning and evening. Power demands, however, can fluctuate throughout the day, and smaller-scale changes in demand can be caused by unpredictable weather damage or random usage.

Power-grid operators in specialized control centers are responsible for monitoring the availability of various generation sources for distribution throughout a grid. That includes the production abilities of both consistently generating nonrenewable sources and more intermittently available renewables, which operators then use to balance supply and demand. That is done through the maintenance of both a baseload of power generation that always matches the minimum power need of the grid and other generation sources that can be quickly turned on or off at any time. Dispatchable sources with slow startup, such as nuclear and coal plants, are thus typically run year-round, and their output is designed to match consistent patterns in demand. Faster-responding dispatchable sources, such as hydroelectric and natural gas, complement the slow sources to meet unexpected high peaks in energy demand and stochastic fluctuations in power usage.

The combination of multiple energy-generation pathways with different strengths allows current power grids to remain as resilient as they are while incorporating new intermittent renewable-energy sources in the face of unexpected disruptions. For all their adaptability, however, grids are not always perfect.

Failure mechanisms and mitigation

The 2021 Texas power grid failed because of a massive drop in electricity generation coupled with a sudden increase in power demand. When cold temperatures froze nonwinterized wind turbines and natural gas pipes, they caused major stress to the power grid. At the same time, the low temperatures led many users to increase their indoor heating, and electricity demand significantly rose to levels typical only during the peak demands of summer. As a result, 46 000 MW of expected power were unavailable while demand rose to 70 000 MW. In response to over half of expected electricity production going offline, the Electric Reliability Council of Texas implemented rolling blackouts by shutting down electricity in neighborhoods for up to 12 hours at a time to reduce demand across the state.[3]

The growing complexity of advanced electrical grid systems means that they have more developed methods of failure prevention, error detection, and built-in redundancy. But the complexity also increases the ways in which an engineered system can malfunction. Source-side faults are caused by an intermittent supply of input resources, such as a cloudy day unexpectedly disabling solar panels or cold temperatures affecting natural gas transportation. They result in power supply discontinuity and are fixed either by the disruption ending or by better strengthening power-generation systems against the elements. Cable faults and deterioration affect power transmission; they are fixed by finding an alternate transmission pathway and repairing the damaged cables. Data communication faults can be caused by hardware, security, equipment, or human error and lead to poor management of available power resources. They can be fixed with high-performance communications networks, including wireless data systems and satellite systems.

Compare those malfunctions with biological modes of failure and interesting similarities emerge. For instance, the heart fails to pump blood adequately when its muscle cells no longer produce enough ATP to keep it beating properly. The process to produce ATP occurs in the mitochondria and is heavily oxygen dependent. A temporary lack of oxygen or accumulated mitochondrial damage reduces the availability of the ATP necessary to keep the heart pumping. That forces alternate pathways to be used instead and triggers the heart to take in more glucose to compensate. Glucose oxidation decreases in proportion to glycolysis increasing, thus letting the heart still produce some, albeit less, ATP without relying on mitochondria or oxygen transport.

Long-term pathway degradation—an inability to draw enough of a resource for distribution—can occur in both biological systems and power grids. In biology, two common causes of heart failure occur when there is not enough energy

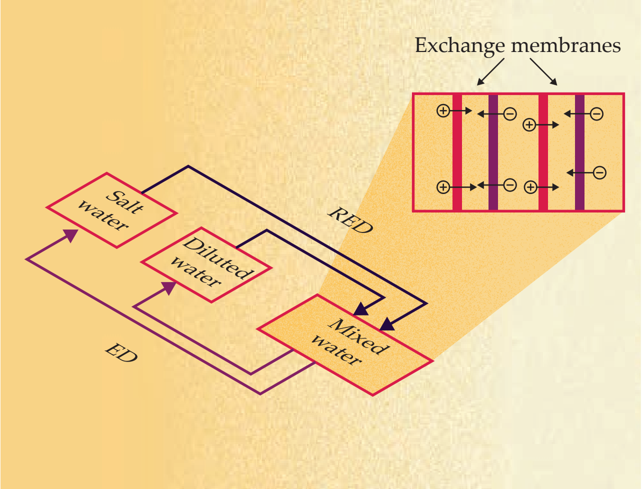


FIGURE 2. CHEMICAL POTENTIAL ENERGY. Two pools of water containing salt water and diluted water are kept separate to store energy. When both are pumped into a third pool and are allowed to mix—a process known as electrodialysis (ED)—charged ions from the salt solution will move past a membrane into the diluted water. An electrical current is generated from the ions passing through membranes. Separating fluids by their salt concentration is known as reverse electrodialysis (RED). (Adapted by Jason Keisling.)

to power cardiomyocytes, cells in the heart responsible for contraction and expansion. The heart either lacks enough strength to pump out blood or doesn't expand sufficiently to fill up with blood. Medications to mitigate the problem target the energy production, not the cells that aren't working properly. Such a response can be paralleled in new power grids by incorporating two types of resolution: active measures, such as creating fast power-generation system restoration, and proactive measures, such as designing weather-resistant hardware.

Resilience and renewable integration

Biological systems and their way of managing energy intake from diverse sources can serve as inspiration for how engineered systems can be optimized to better predict and respond to fluctuating power demand. Humans have a specific sleep–wake cycle governed by the circadian system.[4] Although a body demands energy throughout the entirety of the 24-hour cycle, it has a reduced metabolic need during sleep to avoid overprocessing glucose and having an overabundance of unused ATP.[5] The change in need is accomplished in part by the body's modulating of hormone production to decrease appetite and energy intake and expenditure.[6,7] That the internal system automatically responds to a regular change in supply is one example of evolved biological sensing mechanisms.

The molecular machinery of sensing and responding to fluctuations in nutrient levels also involves modulating pathways that respond to different nutrient sources and directing each cell's ability to meet its energetic needs. For example, the protein glucokinase acts as a sensor in pancreatic cells to detect hyperglycemia. Glucokinase is responsible for initiating a sequence of events that eventually results in insulin secretion to increase cellular glucose absorption.[8]

Those basic principles are paralleled in newly emerging smart grids: A predictive system is capable of detecting changes to power demand via sensors throughout the grid[1] and changing electricity generation in response.[9] Renewable-energy systems that include multiple types of energy generation[10] in particular will benefit from efficient control mechanisms. Developments in that area can help create a fully autonomous smart power-grid system, capable of detecting and responding to changes to all five key elements of an engineered energy grid while helping decrease the need for constant manual oversight.

Current hybrid renewable-energy systems often use a predictive model to optimize the use of multiple renewable sources through regression analysis and Monte Carlo simulation techniques.[11] They also use real-time monitoring and control through advanced networking and information collection, which optimize costs and energy. That coordination among multiple energy types involves using smart detectors to both obtain real-time monitoring and collect data on the status of each source and then prioritize accordingly.

Self-sustainability through dynamic sensing

For a living system to remain self-sustaining, it must be able to manage excess energy. Living systems have evolved to meet that challenge. For example, some fat tissues store unused energy in the form of triacylglycerols (TAGs). A human body contains, on average, one month's worth of energy in TAG storage.

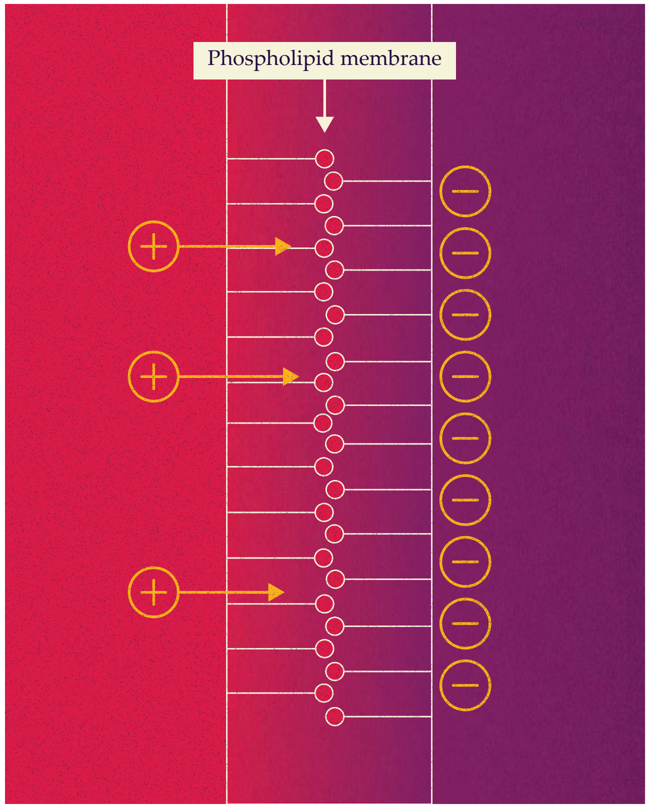


FIGURE 3. CELLULAR CONCENTRATION GRADIENT. A cell wall includes a phospholipid membrane, across which there is an ion gradient. Inside the cell are negative ions. Positive ions from outside the cell will move across the membrane to achieve equilibrium. New batteries for future electrical power grids are designed similarly and will be improved as electrical systems are better able to imitate the biological version. (Adapted by Jason Keisling.)

Yet with a constant stream of new nutrients from getting regular meals, the body usually has little need to actively ration and use those reserves. It is able to dynamically recognize the availability of energy sources and balance the supply with the demand of what cells need to function. When the amount of incoming energy dwindles, however, the pathways that govern TAG storage and metabolism slow down. A decrease in food consumption causes two major metabolic effects with respect to TAG: The body begins rationing the use of stored energy by breaking down TAGs, and it halts energy storage by suppressing the creation of TAGs.[12] Dynamic modulation of stored energy in response to changing energy accumulation is a major success of biological systems.

The need to maintain dynamic sensing in an engineered grid is highlighted by the unexpected shutdown of many northern Spain wind turbines in 2009, when power-grid resilience was aided by energy storage. While wind power production had been projected to supply 45% of total energy demand, the shutdown caused production to fall to a mere 16%. In response, Spain's energy control center immediately increased hydroelectric power and drew on pumped-storage hydropower (PSH) while also starting coal and gas power stations to make up for the sudden drop in supply. Within hours, electricity supply and demand had reached equilibrium again.[13] No outages were reported throughout the duration of the incident, even though the grid experienced a sudden loss of its active power generation.

Note what made that possible—enough stored power in an easily accessible form that could immediately fill power demand when the wind turbines were shut down and multiple ways of generating electricity through human-controllable, dispatchable means. As power-grid systems shift to incorporate decentralized controls and renewable sources, larger portions of power generation will be susceptible to unpredictable external events. To remain self-sufficient, modern grids will need to have a reliable power supply with great energy autonomy, be able to share excess energy with stations experiencing a deficit, and improve energy storage capabilities.

Various forms of stored energy exist in biological systems. They can be immediately converted to ATP or can be kept unused for extended periods of time. Control and feedback mechanisms help in determining what type of energy is stored, how it is stored, and how it is broken down. One of the main forms of energy storage in biological systems is glycogen—a branched molecule of glucose that can be rapidly broken down and converted to energy. Glycogen is mainly formed when living systems do not need to use energy immediately. In addition to glycogen, fats are stored forms of energy that can be converted to usable ATP over the longer term. While glycogen is mainly stored in the liver, fats can be stored throughout the body, creating built-in redundancies.

Energy from gradients

Modern smart grids also have decentralized energy storage for emergencies. The current US power-grid system uses several electrical energy storage systems, such as PSH and advanced battery energy storage. Looking globally, more than 90% of energy storage is via PSH. That type of system uses energy to pump water uphill to higher reservoirs in times when excess energy is available. Then, when electricity is needed, water is released from the higher reservoir through a hydroelectric turbine that produces electricity from kinetic energy.

The PSH method of energy storage uses a topographical gradient to drive energy generation. For large-scale energy storage, it doesn't use traditional batteries, which often incorporate rare and toxic materials.[14] Although it is a promising way of incorporating renewable resources for energy storage, PSH does have its limitations. Namely, it requires specific geographical conditions. Thus, taking advantage of hydropower with the addition of an electrochemical gradient—eliminating the need for specific topographic requirements—can be a worthwhile area of future research. Fortunately, inspiration can be derived from biological systems that employ concentration gradients.

Previous work has investigated batteries that store energy through electrodialysis (ED) and reverse electrodialysis (RED). Both methods utilize the chemical potential difference between two solutions of different salt concentrations.[15] As seen in figure 2, ED separates fluids by their salt concentration to store energy, and RED reverses the process: When the solutions are mixed, charged ions pass through a membrane and their movement generates an electrical current. Previously, RED was mostly used to harness free energy. But recent models have successfully used the chemical phenomenon to create an energy storage system that takes advantage of the chemical potential energy and converts it into electrical energy.

The first installation of an RED-based power plant was in the Netherlands in 2014. Following suit, another was installed in southern Italy, where experimental data on the efficiency of the system was recorded for the first time. The RED prototype, which was tested over a five-month period, was able to endure changes in environmental conditions without experiencing a significant loss in performance.[16]

Looking to the future, improvements in membrane technology will pave the way toward increasing the cost-efficiency of sustainable, RED-based power plants. The principles used in the concentration-gradient flow battery are based on biological systems. As seen in figure 3, cells also use membrane potential and the concentration gradient.

We propose that researchers can continue to enhance electrical power grids by drawing inspiration from biological systems. Improvements such as increasing resilience and flexibility in response to stochasticity, exploring innovative methods of energy storage, and strengthening intercommunication between energy distribution systems are within our reach.

We thank Zeroun Ignis (Harvard College, Cambridge, Massachusetts) for his research assistance on an earlier version of this piece.

REFERENCES

1. E. Y. Song, G. J. FitzPatrick, K. B. Lee, *IEEE Sens. J.* **17**, 7723 (2017).
2. D. J. Thompson, W. C. H. Schoonenberg, A. M. Farid, *IEEE Access* **9**, 68837 (2021).
3. G. Golding, "Texas electrical grid remains vulnerable to extreme weather events," Federal Reserve Bank of Dallas (17 January 2023, updated 24 January 2023).
4. M. J. Aminoff, F. Boller, D. F. Swaab, in *Handbook of Clinical Neurology,* vol. 98, P. Montagna, S. Chokroverty, eds., Elsevier (2011), p. vii.
5. C. M. Jung et al., *J. Physiol.* **589**, 235 (2011).
6. S. Dornbush, "Physiology, Leptin," N. R. Aeddula, ed., StatPearls (updated 10 April 2023).
7. G. Pradhan, S. L. Samson, Y. Sun, *Curr. Opin. Clin. Nutr. Metab. Care* **16**, 619 (2013).
8. F. M. Matschinsky, D. F. Wilson, *Front. Physiol.* **10**, 148 (2019).
9. For information on the US Department of Energy's innovations, see www.energy.gov/electricity-insights.
10. K. S. Krishna, K. S. Kumar, *Renew. Sustain. Energy Rev.* **52**, 907 (2015).
11. M. K. Deshmukh, S. S. Deshmukh, *Renew. Sustain. Energy Rev.* **12**, 235 (2008).
12. S. Kersten, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1868**, 159262 (2023).
13. C. Goodall, "Spain's variable wind and stable electricity networks," *Carbon Commentary* (15 November 2009).
14. W. J. van Egmond et al., *J. Power Sources* **340**, 71 (2017).
15. R. S. Kingsbury, K. Chu, O. Coronell, *J. Membr. Sci.* **495**, 502 (2015).
16. M. Tedesco et al., *J. Membr. Sci.* **500**, 33 (2016).

PT

FIXING THE PHD QUALIFYING EXAM

Timothy DelSole and Paul A. Dirmeyer

FIZKES/SHUTTERSTOCK.COM

Tim DelSole (tdelsole@gmu.edu) and **Paul Dirmeyer** are professors of climate dynamics at George Mason University and senior research scientists at its Center for Ocean-Land-Atmosphere Studies in Fairfax, Virginia. Together they have overseen the PhD qualifying exam in their department for more than two decades.

For the past five years, the faculty in the department of atmospheric, oceanic, and Earth sciences at George Mason University has used a qualifying process that overcomes many of the shortcomings of traditional exams.

As senior scientists, we have navigated the challenging waters of the PhD qualifying exam—both as students taking it and as professors administering it. As students, both of us excelled academically, yet we anticipated the qualifying exam with anxiety and dread. How could we not? The professors judging us could inquire about any aspect on which they were expert. We prepared for the exam by revisiting our coursework, aware that even the most thorough review might not suffice, as some professors saw the exam as an opportunity to push students beyond core knowledge.

We recall preparing diligently for specific topics about which we were never queried, and thus we were unable to showcase our extensive preparation. We recall knowing the answers to some exam questions in retrospect, but in the pressure of the moment, we couldn't remember them. What's more, passing the qualifying exam left us no closer to defining our thesis research direction.

Later we discovered that professors also approach the qualifying exam with anxiety and dread. The consequences of the exam put immense pressure on professors to craft questions that can accurately gauge a student's potential. The exam's duration—mere hours or days—seems inadequate for making such a significant judgment on a student's future. Students commonly stumble over questions, compelling us to look past their mistakes to infer their potential. Such a process depends heavily on subjective judgment. And the fact that those judgments usually rest with a few faculty members raises concerns about fairness and the inclusion of diverse viewpoints. Even more troubling, insights gained from the exam are often minimal; we could usually predict a student's outcome from their past coursework performance.

When questioning the rationale behind the qualifying process, we find that the arguments for retaining the traditional qualifying exam's written and oral components do not hold up to scrutiny. One commonly stated goal is to assess the student's grasp of the core knowledge attained from one to two years of coursework. We wondered whether failure to pass the qualifying exam reveals more about the inadequacies of the courses than those of the student. If the true goal is to ensure mastery of core knowledge, a simple way to ensure it is to make that mastery the standard for passing the courses.

Another common argument is that the exam tests the student's ability to synthesize concepts across disciplines. Although synthesis is a valuable skill, so are deep dives—Nobel Prizes, for instance, are often awarded to scientists who relentlessly pursue a narrow scientific question.

Yet another argument is that the exam assesses creativity. A student's creative strengths, however, may lie in emerging fields, such as artificial intelligence, that are not covered by the traditional written and oral components and that may fall beyond the expertise of the examiners.

There also are compelling arguments against the traditional format. The qualifying exam is the most stressful milestone in a graduate degree. Not all students perform well under stress, regardless of their capacity for deep thought. Moreover, the traditional exam embodies a contradiction: It is contrived and fails to mirror the realities of actual research. A successful scientific career relies on conducting actual research, not on taking exams. Judging a researcher's potential by testing them on tasks unrelated to their future work is inherently flawed.

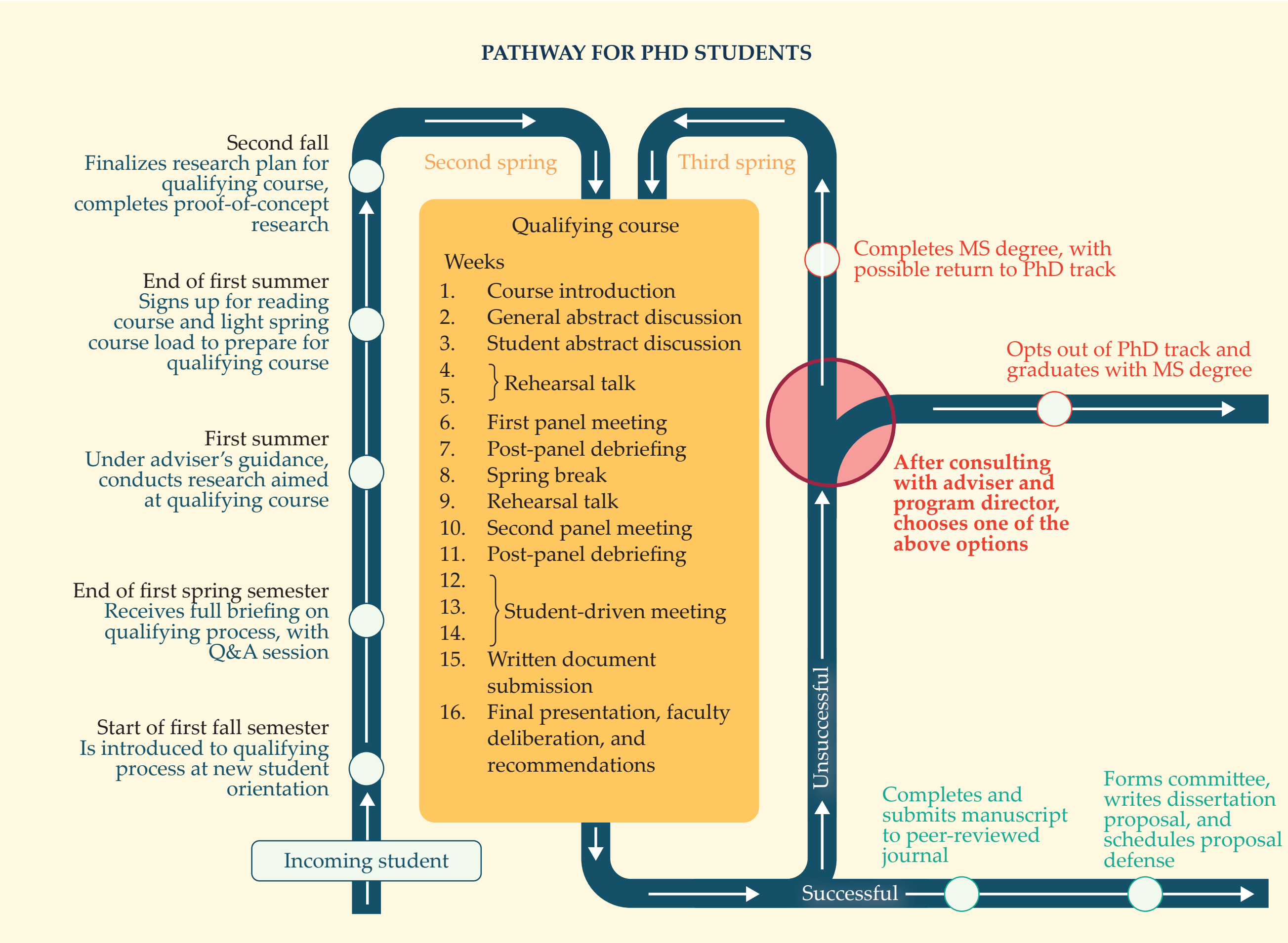


FIGURE 1. THE QUALIFYING EXAM, REVISITED. An incoming PhD student in the atmospheric, oceanic, and Earth sciences department at George Mason University completes a series of milestones (blue) in the semesters before the qualifying course. The milestones of the qualifying course are listed in the central box. The successful student goes on to submit a manuscript to a journal and form a dissertation committee (green). The unsuccessful student may take the course again or exit to the master's degree track (red).

Despite those shortcomings, there remains a compelling necessity for qualifying exams: Experience shows that some students, despite passing their courses, struggle to complete a dissertation within the typical five-year doctoral program. Identifying those students early allows all parties to move forward without investing years of effort into a PhD journey that may ultimately be unsuccessful.

Recognizing the limitations of the traditional qualifying exam, we convinced our colleagues in the department of atmospheric, oceanic, and Earth sciences at George Mason University to scrap the traditional format five years ago and institute a new process that was designed to overcome its shortcomings.

A different approach

Figure 1 provides a visual summary of the new process and its place within the overall PhD track. It centers on a semester-long course typically taken in the spring of a student's second year. In the course, the student works with their adviser to formulate a project that will lead to a research paper.

The student presents their paper idea in two meetings. In the first, the student delivers their proposal during the first 15 minutes, followed by a 75-minute period during which panel members pose critical questions about it, as shown in figure 2. Half the department's faculty members are present in that meeting. About four weeks later, the student presents to the other half of the faculty. The two faculty panels function independently without communicating with the other. The autonomy provides the student with two independent opportunities to present their work at their best, free from biases influenced by the previous performance.

After each meeting, the student receives written feedback from each panel member on various aspects of their paper idea. That feedback includes an assessment of the student's grasp of the relevant literature, their physical understanding of the scientific problem, their ability to perform quantitative analysis, and their effectiveness as a communicator. Panel members evaluate each of the categories, but they don't assign a grade. The feedback is intended to be constructive—to help the student identify areas in which they need improvement.

By the end of the semester, the student submits either a manuscript that is nearly ready for submission to a peer-reviewed journal or, if the research is not yet completed, a proposal for a scientific paper that incorporates their original research. All panel members read the student's document. On the final day of the course, the student gives a longer oral presentation to the entire faculty. The faculty members then discuss and make recommendations to the program director.

A better way to evaluate students

The new qualifying process has several advantages over the traditional format. First, instead of assessing a student's knowledge, faculty members evaluate the student's ability to perform the activities critical to scientific inquiry: identifying a scientific problem, devising solutions, and engaging in discourse. Second, the process spans an entire semester, so decisions on a student's performance are not based on a singular moment. Third, each student chooses their own research topic, affording them the opportunity to showcase their creativity.

Furthermore, a student receives questions tailored to their chosen topic. That focus avoids the pitfall of evaluating them on their response to questions far removed from their prepared area. Although confronting a student with unanticipated questions can have merits, using that approach as a sole determinant of the student's future is risky. A semester-long engagement provides a more reliable assessment of potential.

The new qualifying process also offers each student multiple opportunities to succeed. Because the two faculty panels are independent, a student who struggles in the first meeting can present a revised version of their work to a fresh audience. Some students might face special challenges in the oral presentation, such as stage fright or language deficiencies. To address that issue, the new process includes a written submission as an additional means to demonstrate the student's abilities.

Grading in the qualifying process does not rely on averaging individual scores. The primary goal is to identify evidence of the student's research capabilities. The evidence may not be uniformly apparent across all components, but outstanding performance in a single aspect can eclipse weaker performance in other areas. Additionally, we consider the student's progress throughout the semester because improvement is often a key indicator of potential success in research.

Another advantage is that a student has ample opportunity to revise their work. Even the best student may not fully explore their ideas initially. Our semester-long process mimics the peer-review process and typically exposes any serious shortcoming that may exist in a student's proposal. Observing how students adapt to constructive feedback often provides more insight into their potential than does reviewing their initial proposals.

Unavoidably, subjective judgments affect the final decision, and they may be influenced by biases tied to race, gender, sexual orientation, or disability. Even the traditional format, with its fact-based questions, involves subjective judgments in deciding the acceptability of a student's response. One way to counter biases is to involve a diverse panel of judges whenever subjective judgments come into play. In the traditional format, the decision often rests with a select few faculty members. In the new format, the entire faculty openly participate in the decision-making process, which brings a wider range of perspectives into the discussion.

By distributing responsibility across all faculty members, the new process also lightens the burden on individual advisers, who often hesitate to single out their own struggling students. When a student is redirected, their adviser usually appreciates the collective intervention.

FIGURE 2. A STUDENT in the new PhD qualifying process presents a proposal for a paper to half the department's faculty. About four weeks later, the student gives a revised presentation to the other half of the faculty. The two faculty groups function independently without communicating with the other.

Although the new format requires a greater investment of time from faculty, productive scientists are accustomed to allocating time for conferences and peer-review duties. And the new qualifying process calls for minimal preparation by faculty, with only modest tasks required post-meeting, such as filling out evaluation forms. When it comes to peer-review services, the question is how to best manage one's time reviewing others' work. Allocating a portion of that time to assisting students in one's own department proves to be a sound investment in upholding the quality and integrity of the qualifying process. Ultimately, the efforts produce better student outcomes, which, in turn, cast a positive light on the faculty and the department.

Fellow scientists who hear about our qualifying process are often doubtful about its feasibility in their own departments. They cite factors such as a large student population. We are confident, however, that the new process can be tailored to any department. Our PhD program at George Mason has a dozen faculty members and admits three to six candidates per year. For larger departments, splitting students and faculty into smaller cohorts operating in parallel is a feasible solution.

Another concern has been the perceived inefficiency of involving faculty who lack expertise in a student's chosen topic. But we have found the opposite to be true: Observing how the student articulates their research to nonspecialists, who nonetheless possess broad scientific knowledge, has several advantages. Incorporating diverse expertise in faculty panels, for instance, ensures that a mix of technical and foundational questions will be addressed, which makes the evaluation more thorough.

The new process also encourages faculty to engage with each student out of genuine interest, thus fostering a less adversarial interaction than the traditional approach. The reversal of the conventional roles of teacher and student mirrors what a student will encounter in advanced doctoral research. Furthermore, the shift in dynamic creates opportunities for a student to demonstrate creativity in handling conflicting criticisms that arise from reviewers with different knowledge backgrounds.

One issue that has generated considerable debate among our faculty is the grading policy. Currently, a student who passes the qualifying course receives either an A or a B. The A

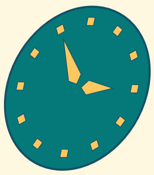
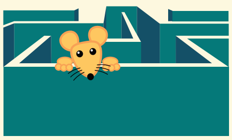
	SHORTCOMINGS OF THE TRADITIONAL APPROACH	ADVANTAGES OF THE NEW QUALIFYING PROCESS
	• The focus is on the student's existing knowledge.	• The focus is on the student's ability to identify an open science problem, communicate a viable plan, and demonstrate their capability to conduct research.
	• A timed exam doesn't test the student's capability to conduct independent research.	• A 16-week course allows the student ample time for contemplation, review, response, and revision.
	• The exam environment is artificial and not like the settings where the student will present their science.	• Papers and proposals are elements of the profession. • Short presentations are like those given at conferences and workshops.
	• The exam setting magnifies stress, and the agenda is outside the student's control.	• The student selects their topic before writing their abstract and panel presentation. • The student has time to revise their ideas and multiple opportunities to present them.
	• The final decision is made by a few select faculty members.	• All department faculty members participate in the evaluation process and contribute to the final decision for each student.
	• The exam does not help the student develop research and communication skills or complete a publishable manuscript.	• One written and two oral elements allow the student to develop their skills. • The process jump-starts the student's submission of their first paper.
	• Many faculty members have little exposure to the student's research ideas until late in their PhD program.	• Faculty members learn of the student's topic early and can discuss it with them. • The student can recruit interested faculty to their dissertation committee.

FIGURE 3. COMPARING QUALIFYING PROCESSES. The traditional PhD qualifying exam assesses a student's knowledge through written or oral exams that last a few hours or days. The new PhD qualifying process evaluates a student's progress toward a publishable research paper over a 16-week semester.

grade, however, is reserved for students who submit a manuscript that the faculty believes can be refined into a publishable paper after a few months of revision. That's a high standard, and not all exceptional students meet it.

We believe that a significant distinction exists between a student who develops a nearly publishable paper in their second year and one who does not, and the grade assigned to each one is intended to reflect and reward that difference. Moreover, the standard is attainable: One or more students achieve it each year.

Support for students

Most second-year graduate students find the prospect of formulating and defending a publishable scientific analysis in a single semester daunting. Indeed, many students have never presented their own research in front of a group of scientists. To address the issue, we have implemented support mechanisms to assist each student throughout the qualifying process.

First, the student works with their adviser to formulate an idea that will be integrated into their dissertation. If the student is supported by a grant, they are encouraged to select a topic related to that grant, but their contributions must be independent. The new format provides opportunities for the student to innovate while still benefiting from their adviser's guidance.

Advisers must avoid overdoing their guidance; otherwise, the process becomes an evaluation of the adviser instead of the student. Our tenet is that the process should not disrupt the natural interaction between student and adviser. Reasonable guidance includes suggesting research topics, offering feedback on presentations and written materials, assisting in problem diagnosis, and helping the student devise strategies for solutions.

Beyond that, it is left to the student to use the information they receive. The adviser should avoid writing code on behalf of the student or producing text that could be copied into the student's written submission. The student is expected to defend their ideas without assistance from their adviser. In addition, we advise each student to reduce their course load or take a reading course during the qualifying process to allow more time for conducting independent research.

The two of us currently lead the qualifying course, guiding students through the process. We listen to practice talks prior to panel meetings and offer guidance on delivering effective presentations. Students consistently underestimate the level of detail necessary to communicate their research plan effectively, and some are unwittingly too dependent on their advisers to address basic questions related to their project. We strive to inspire students to take ownership of their work and to thoroughly understand the models and data that they use. Practice talks can expose potential research flaws early enough for students to make adjustments before their first panel meeting. For a point-by-point comparison of the traditional exam's shortcomings and the advantages of the new process, see figure 3.

We also meet with each student after their panel meeting to discuss written feedback from the faculty. That feedback resembles the kind that might be encountered during a genuine peer-review process, encompassing not only constructive feedback but also potential contradictions. As educators, we believe that it is crucial not to shield students from that reality. Instead, we strive to expose students to diverse perspectives and help them interpret the resulting feedback constructively.

Sixteen weeks is hardly enough time to complete a serious research project. Accordingly, we encourage every PhD student to begin preparing for the qualifying process as soon as they enter our program. Any research conducted by a student during their time in the graduate program is permitted for use in the qualifying process. And the qualifying process is explained in detail to PhD students at the end of their first spring semester. The early introduction helps instill a productive mindset in each student as they approach their first summer in the program—a period devoid of course distractions that allows them to focus wholeheartedly on their research goals.

Teaching research and communication

As advisers, we try to strike a balance between providing assistance and giving students space to develop their own thinking. To do that effectively, we've adopted a strategy inspired by what's known as the Heilmeier catechism. George Heilmeier, who led the Defense Advanced Research Projects Agency in the mid 1970s, crafted a series of questions that every good proposal should answer.[1] The list distills years of wisdom into a concise question set. We have adapted it to create the following questions that every research project should address:

►What are you trying to do? State your objectives without jargon.

►Who should care? If you are successful, what difference will it make?

►What research has been done about the topic in the past?

►What is the precise gap that you are trying to fill?

►What is new in your approach?

►Why do you think your approach will be successful, and how will you measure success?

The simplicity of the questions can deceive students into underestimating the effort needed to answer them effectively. To ensure the development of thorough answers, students write an abstract for their projects early in the semester. Those abstracts are then shared with the class, sparking discussions about best practices and common pitfalls when communicating scientific ideas. Invariably, some abstracts fail to address one or more key questions. Some students are convinced that they have responded adequately to a question until further discussion reveals gaps in their explanation. The value of carefully addressing the questions often hits home during the discussions.

Many advanced students begin the semester with a well-defined paper plan but are surprised by the challenge of communicating their plan to others. The situation reflects a fundamental reality: Success as a scientist relies on communication as well as critical thinking. Proficient communication skills are vital for today's PhD graduates, whether it's for securing funds, responding to peer review, teaching, or mentoring. Although research and communication skills are often regarded as distinct, we find much truth in the adage that poor communication may reflect poor thinking.

The decision

The final decision of whether a student passes the qualifying process is a collective one. Although faculty opinions may vary initially, discussions usually lead to a consensus. Students who do not pass typically display one or more common traits: an inability to demonstrate quantitative analysis; a failure to understand the relevant literature; insufficient familiarity with the data or model chosen for study; an inability to answer questions related to calculations, models, or assumptions; an inability to articulate how the proposed research addresses key questions; difficulty in communicating the research plan; and an inability to understand or convey the relevant concepts.

A student at risk of failing is typically alerted after the panel meetings. That early notification gives them time to make improvements and address the concerns. Consequently, a negative outcome is rarely a surprise to the student. Some students have voluntarily withdrawn from the qualifying process during the semester after recognizing that they were unlikely to meet the necessary requirements. The self-selection process allows students to make informed choices about their academic path and potentially explore alternative options that better align with their capabilities and interests.

A student who fails the qualifying process has the opportunity to reframe their work into a master's thesis and complete that degree instead. Some students have retaken the PhD qualifying process and progressed to candidacy, having benefited from the early identification of areas for improvement.

A fortunate byproduct of the new qualifying process is that it energizes the student for their dissertation research. Throughout the semester, the student has multiple opportunities to present and refine their research ideas. When the student successfully passes the qualifying process, they do so with confidence in their ideas and typically reduce the time required to complete their first paper. Compare that with the traditional qualifying exam, which a student might pass without gaining any clearer direction for their research. Another byproduct is that faculty members learn of the student's research topic early and may develop productive dialogues with them. Likewise, a student becomes familiar with faculty interests early on. That helps them identify suitable candidates for members of their dissertation committee.

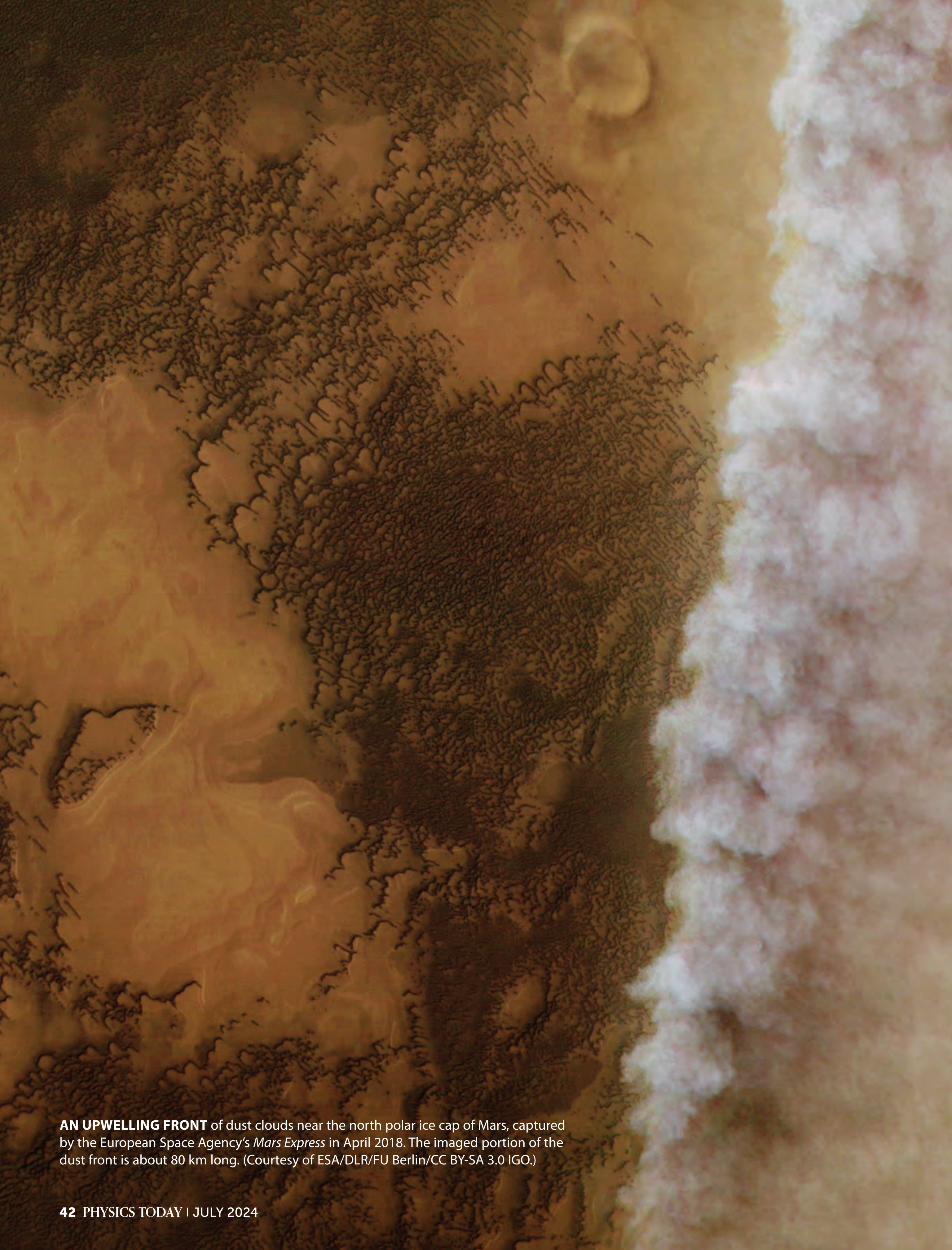
Each year we watch students rise to the challenge of crafting original ideas. Witnessing students' growth and maturation is inspirational. Indeed, the faculty also learn from the process. Many of us feel that our skills as research advisers improve as a result of it. The new format elevates the qualifying process from a routine student assessment to a shared journey of scientific discovery.

We thank our fellow faculty members for their patience and for helping to refine the qualifying process, testing it (with two of our own students), and ultimately integrating it into the department's curriculum. We also thank faculty and students for their feedback on the manuscript.

REFERENCE

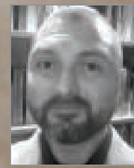
1. G. H. Heilmeier, "Some reflections on innovation and invention," *Bridge*, Winter 1992, p. 12.

PT



AN UPWELLING FRONT of dust clouds near the north polar ice cap of Mars, captured by the European Space Agency's *Mars Express* in April 2018. The imaged portion of the dust front is about 80 km long. (Courtesy of ESA/DLR/FU Berlin/CC BY-SA 3.0 IGO.)

Erdal Yiğit is a professor of physics at George Mason University in Fairfax, Virginia. His research focuses on the global modeling and observation of internal-wave effects in planetary atmospheres.



Exploring Mars's harsh atmosphere

Erdal Yiğit

Getting humans to Mars is difficult enough. But things won't be any easier after they arrive: The red planet's climate and weather are anything but friendly.

With its reddish-orange color in the night sky, Mars awakened the curiosity of early astronomers. The ancient Egyptians called the planet Her Desher, meaning “the red one.” Because of the color, ancient civilizations associated the planet with their god of war: the Babylonians called it Nergal; the Greeks, Ares; and the Romans, Mars. After the invention of the telescope in the beginning of the 17th century, the red planet was observed in more detail. Galileo Galilei made the first telescopic observation of the planet in 1610. By 1636 the Italian astronomer Francesco Fontana had made a simplified sketch of the Martian disk.

The creation of a detailed surface map of Mars by Giovanni Schiaparelli in 1877 was a crucial step in the planet’s exploration. He termed a series of linear features *canali*, which was unfortunately mistranslated into English as “canals,” not “channels.” In the late 19th century, that error famously led to speculation that they were artificial waterways built by intelligent aliens. Today we owe our most accurate map of Mars to the Mars Orbiter Laser Altimeter (MOLA), an instrument on the *Mars Global Surveyor* spacecraft that mapped Mars’s surface from 1997 to 2001.

Figure 1, comparing Schiaparelli’s map with the MOLA map, demonstrates the extent of development in characterization of Martian surface features. Because Mars lacks oceans and thus has no sea level, elevation on the red planet is measured relative to a reference surface level known as the Mars datum, or areoid: an equipotential surface whose mean value at the equator is equal to the red planet’s mean radius. The planet’s topographical extremes include the largest volcano in the solar system, Olympus Mons, which rises about 25 km above the surface; and one of the solar system’s vastest impact craters, Hellas Planitia, which descends about 7 km below the surface.

It takes about nine months for a spacecraft to reach Mars. Since the pioneering *Mariner 4* flyby in 1965, the red planet has been the destination of numerous orbiters, landers, and rovers. Those missions have studied the planet’s geology, atmosphere, and weather and the question of whether it ever harbored life. Seven orbiters and three rovers are currently exploring the Martian atmosphere and surface. NASA operates two of the rovers, *Curiosity* and *Perseverance*, which are situated near the equator. *Curiosity*’s main goal is to assess whether Mars ever had an environment capable of supporting life; *Perseverance*’s objectives are to look for signs of ancient life and to collect samples that are intended to be retrieved by a future mission. The other rover, *Zhurong*, is the first Chinese vehicle on Mars. It was deployed in 2021 by the Tianwen-1 lander.

The rovers are mobile laboratories and observatories: Traveling at approximately 0.1 km/h, they carry multiple instruments, take Martian “biopsies,” and perform remote sensing. The data gathered have revealed much about the red planet’s surface, atmosphere, and ionosphere. They are also informing our understanding of how space weather in the solar system affects Mars. Space agencies are using the results to help prepare for human exploration of Mars in the not-too-distant future.

The Martian atmosphere

With an axial tilt of 25.2°, which is relatively close to Earth’s 23.4°, Mars also has four seasons, but each is about twice as long as Earth’s. Moreover, the seasonal differences are drastic because of Mars’s orbit, which is far more eccentric than Earth’s. The red planet has strong seasonal and interhemispheric differences in atmospheric waves, air circulation, chemical composition, dust distribution and heating, and weather.[1]

Mars’s surface and atmosphere are drastically different from Earth’s. The red planet’s surface is mainly covered by fine dust, sand, and rocks and boulders of varying sizes, which are collectively referred to as Martian regolith. Figure 2 shows the composition of the surface as viewed by *Perseverance* and *Curiosity*. On average, Martian soil is composed of 98% mineral matter and only 2% air and water. In comparison, Earth’s soil contains about 45% mineral matter and 50% air and water. The

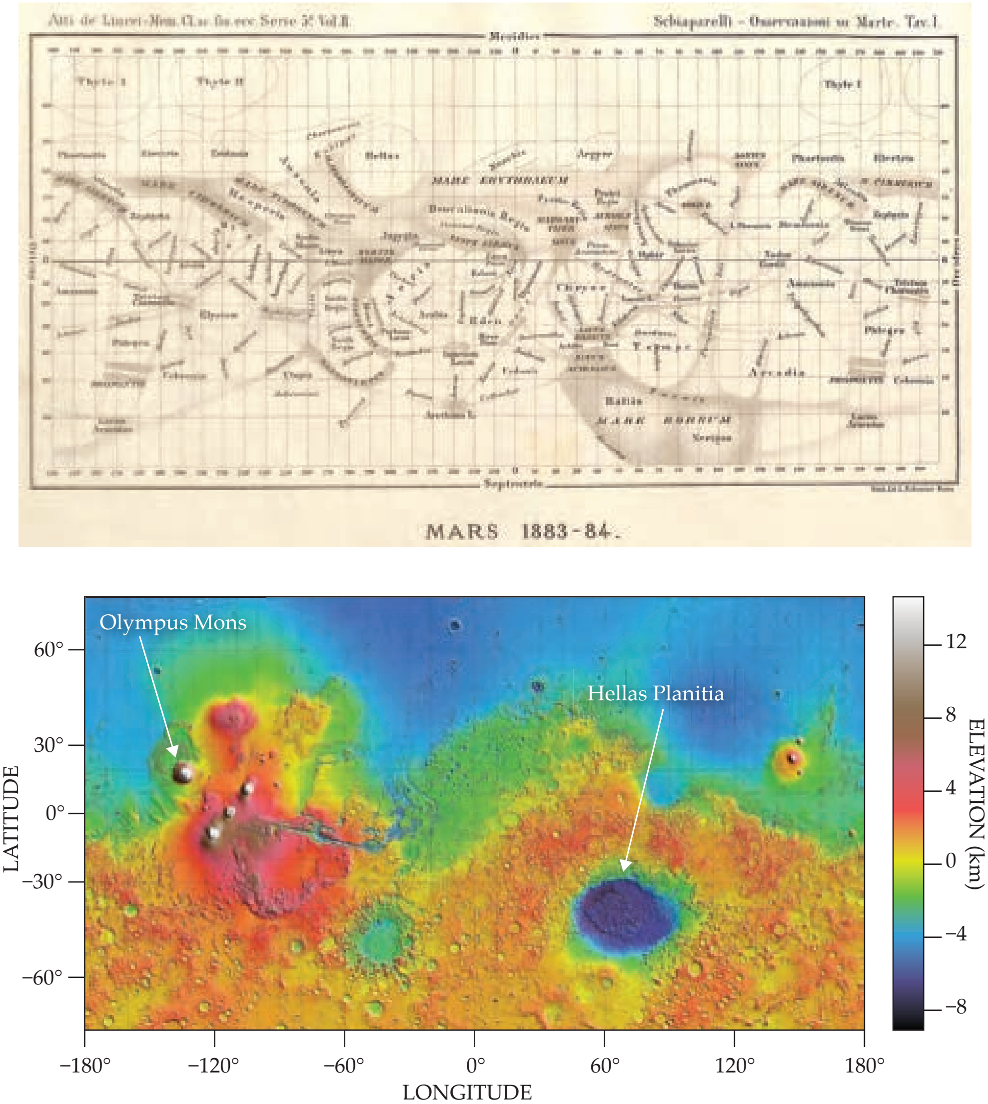


FIGURE 1. MAPS OF MARS, old and new. At top is an 1896 map of the Martian surface created by the Italian astronomer Giovanni Schiaparelli on the basis of observations he took in 1883–84. Because 19th-century telescopes produced inverted images, the map is flipped: The southern hemisphere of Mars is at the top. At bottom is the most complete topographic map of Mars available today. Generated in 2000, it is based on four and a half years of measurements taken by the Mars Orbiter Laser Altimeter, an instrument onboard the *Mars Global Surveyor*. Higher elevations are colored red to white; lower elevations, blue to black. The red planet's two topographical extremes, Olympus Mons and Hellas Planitia, are marked on the map. (Top image courtesy of Giovanni Schiaparelli/public domain; bottom image adapted from the MOLA Science Team.)

other 5% is made up of organic matter, which according to our current state of knowledge, is mostly absent on Mars.

Martian soil is essentially analogous to the type found on Earth near iron-rich volcanoes, and it is probably drier than that found in the Sahara Desert. Although it is widely accepted that present-day Martian soil cannot support life, the search for past biosignatures, such as fossils, remains ongoing. Bringing Martian samples to Earth would be a groundbreaking step in that research, but it would be expensive and carries the danger of cross contamination. Moreover, that research has historically been complicated by the presence of structures in Martian soil that resemble fossils yet are not.

More than 95% of the Martian atmosphere is composed of carbon dioxide. With a surface pressure that's 0.6% of Earth's, Mars's atmosphere is thin, primarily because of its weak gravity: Mars is approximately half the size of Earth, and its gravity is about 38% of ours. Because the escape velocity of air is determined by both a planet's gravitational field and its radius, atmospheric species can more easily escape to space from Mars than from Earth.

Mars's atmosphere extends about 250 km above the surface. It is divided into three distinct layers, which are determined by altitudinal changes in atmospheric temperature.[2] The first layer is the troposphere, which extends from the surface to 50 km above the Mars datum. Next comes the mesosphere, which ranges from 50 to 120 km above the surface, and above that is the thermosphere, which extends to about 250 km. Unlike Earth, Mars has no stratosphere: Because its lower atmosphere does not contain ozone, it does not absorb solar UV radiation like Earth's and therefore lacks the temperature inversion that causes that layer to form.

Above the thermosphere is the exosphere, where the atmosphere becomes increasingly collisionless—the mean free path of the molecules becomes longer and the collisions between particles become less frequent. The highest portion of the thermosphere is often termed the exobase. Its altitude varies between 210 and 250 km, depending on the solar local time and season, as recently observed by the *MAVEN* (*Mars Atmosphere and Volatile Evolution*) spacecraft.[3]

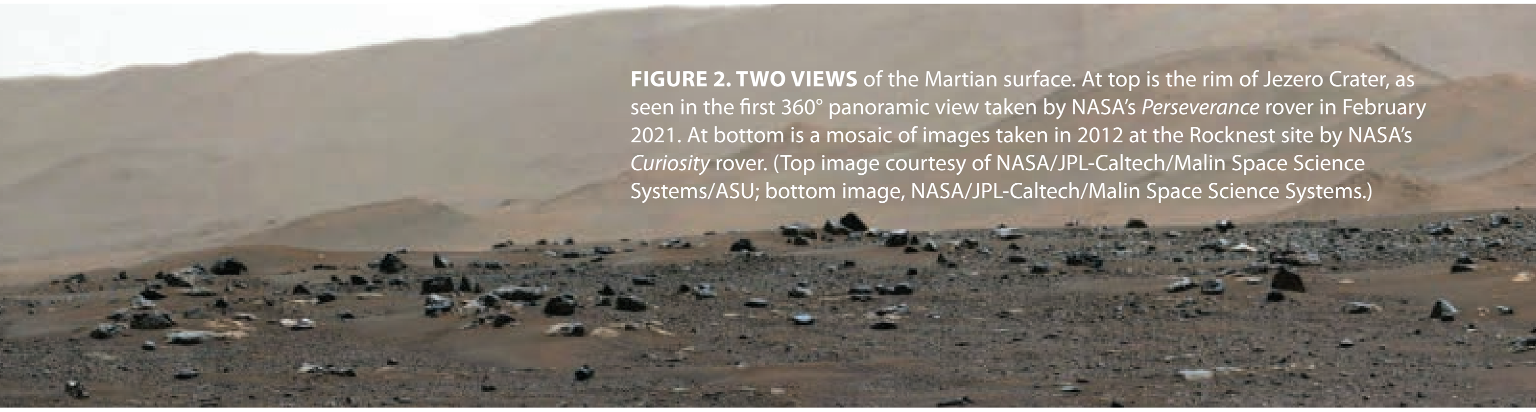
The Martian atmosphere is stably stratified, and buoyancy forces cause lighter air layers to float above the denser ones. Consequently, the air density drops off roughly exponentially with height. The stability of an air parcel can be disturbed by external influences. For example, surface heating initiates convection, which causes lighter, warmer air to rise upward and cool until it reaches its equilibrium level. That convective motion, or

mixing, can occur only under unstable atmospheric conditions, in which lighter air temporarily lies below denser air.

Convection plays a central role in the transport of energy and momentum in planetary and even stellar atmospheres. It can generate atmospheric gravity waves with fast phase speeds. Gravity waves are natural buoyancy oscillations found in all stably stratified planetary atmospheres. By propagating over long distances and transferring energy and momentum from the lower regions of the atmosphere to higher altitudes, they significantly shape the temperature, circulation, and composition of the middle and upper atmosphere.

When atoms and molecules in a planetary atmosphere absorb solar energy, they can become excited, which means that the electrons are raised to a higher energy state than their ground state. If the absorbed energy exceeds the ionization energy of a given species, electrons can be ejected, resulting in the formation of a positively charged—or ionized—atom or molecule and a free electron. The region of ions and free electrons formed by the absorption of solar radiation is the ionosphere, a naturally occurring plasma environment that coexists with the neutral atmosphere.

Mars's ionosphere begins approximately 90 km above the areoid and extends to a varying altitude of about 400–1200 km. That means it coincides with the upper mesosphere and thermosphere. Martian ions are produced by the ionization of carbon dioxide and oxygen atoms, whose lowest ionization energies are 13.78 eV and 13.61 eV, respectively. *MAVEN* observations,[4] however, demonstrate that molecular oxygen ions are a major species up to about 350 km. They are produced by two processes: charge transfer and atom–ion exchange between atomic oxygen and CO_2^+. Above that altitude, atomic oxygen ions become the most abundant ionospheric species. At lower altitudes, the ionosphere interacts significantly with the thermosphere, and the ions exchange energy and momentum with neutral particles and atmospheric waves. The topside ionosphere, on the other hand, is highly variable in density and structure.[5]



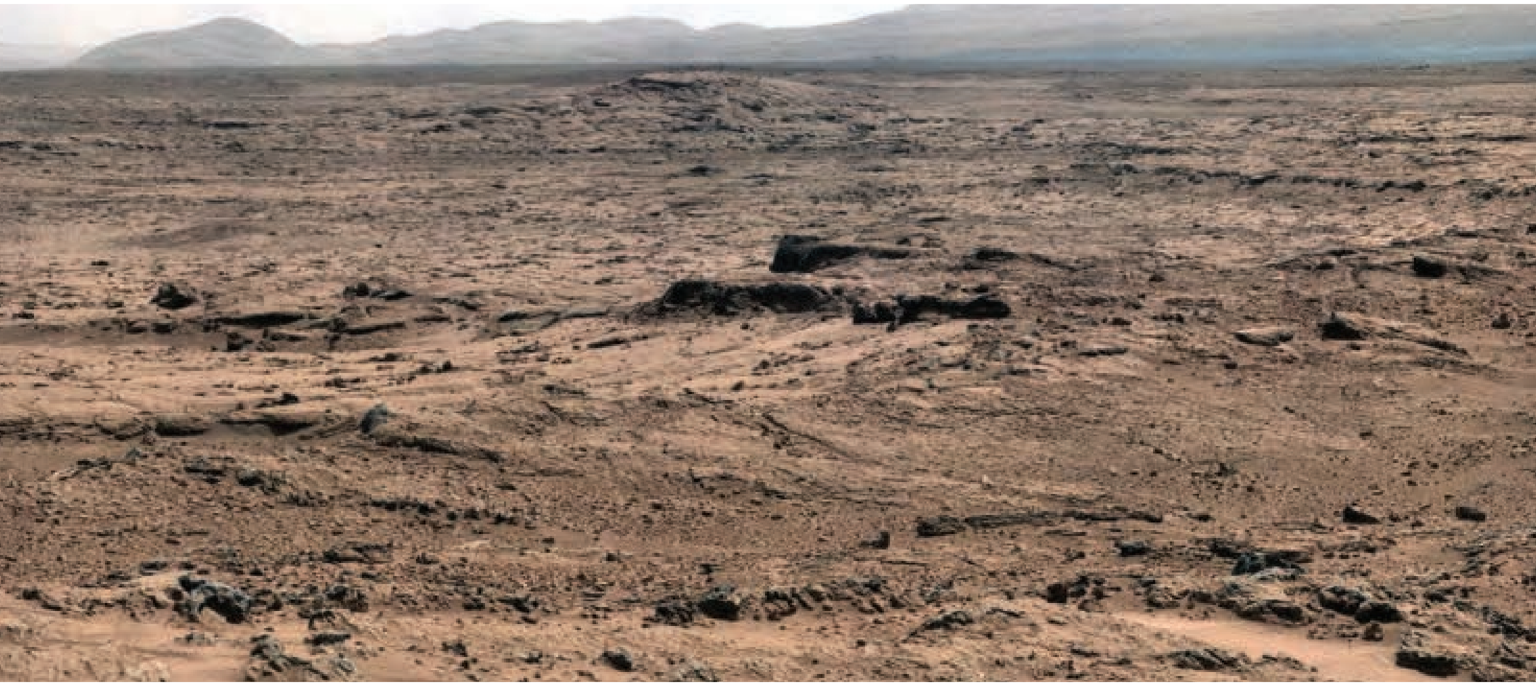


FIGURE 2. TWO VIEWS of the Martian surface. At top is the rim of Jezero Crater, as seen in the first 360° panoramic view taken by NASA's *Perseverance* rover in February 2021. At bottom is a mosaic of images taken in 2012 at the Rocknest site by NASA's *Curiosity* rover. (Top image courtesy of NASA/JPL-Caltech/Malin Space Science Systems/ASU; bottom image, NASA/JPL-Caltech/Malin Space Science Systems.)

Recent *MAVEN* observations demonstrate that the Martian ionosphere exhibits strong local time and seasonal variations with distinct hemispheric asymmetry.[6] It is also responsive to solar wind and solar activity. But ionospheric irregularities of various spatiotemporal scales are also often observed, which may be indicative of upward-propagating gravity waves.

Martian meteorology

Mars's weather is characterized by features familiar to us on Earth, including atmospheric waves, winds, dust, clouds, and water vapor. Like terrestrial climate scientists do for Earth's atmosphere, planetary scientists use numerical models to reveal how different Martian atmospheric layers and weather phenomena interact with each other and respond to external processes. Those models quantify energy, mass, and momentum transfer and are used to predict the future state of the Martian atmosphere on the basis of its current state.[1,7] But because we know far less about Mars than we do about Earth, models of the red planet's atmosphere and weather are less robust than their counterparts on our planet: Present-day terrestrial weather models, for example, continuously assimilate global observational data at a high spatial resolution, but analogous models for Martian weather have not yet been developed.[8]

Gravity waves contribute significantly to Martian weather. Sometimes called buoyancy waves, they are the most frequently observed type of wave on Mars and are produced by meteorological processes in the troposphere.[9] They owe their presence to the counteracting forces of gravity and buoyancy on vertical displacements of air. They are produced primarily in the troposphere by convection, nonlinear instabilities, and atmospheric flow over Martian surface topography. Because of the conservation of energy, gravity waves' amplitudes increase as they propagate upward. In the mesosphere and thermosphere, the amplitudes become so large that the atmosphere cannot sustain the associated fluctuations. Like ocean waves

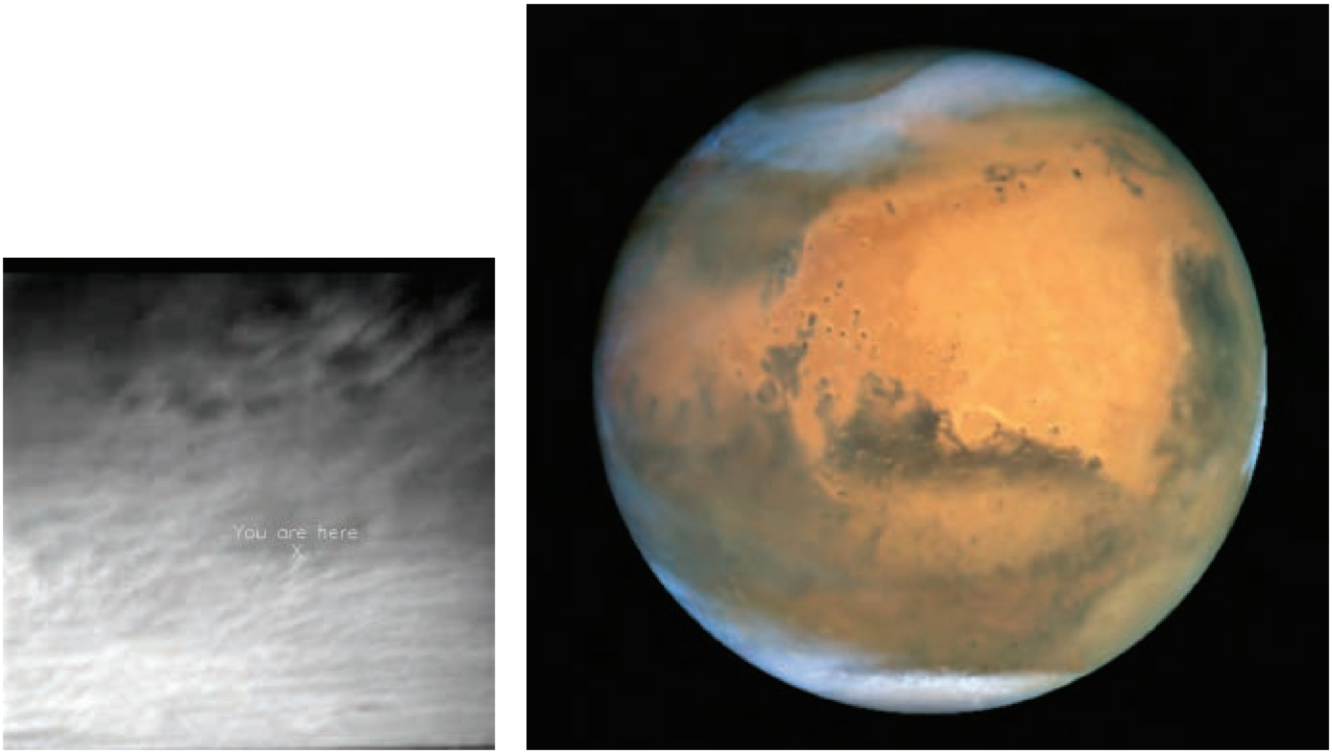


FIGURE 3. EARLY IMAGES of Martian clouds. At left is the first image of clouds on Mars. Taken by the *Mars Pathfinder* lander in 1997, it depicts stratus water-ice clouds at an altitude of about 16 km. The "You are here" sign marks Earth's position in the Martian sky at the time the image was taken. At right is an image of Mars captured by the *Hubble Space Telescope* in June 2001. Water-ice clouds exhibiting wavelike behavior are visible at high latitudes in the northern hemisphere. Seasonal dust activity can be seen near Hellas Planitia in the southern hemisphere at bottom right. (Left image courtesy of NASA/JPL/University of Arizona; right image, NASA/Hubble Heritage Team.)

approaching the shore, the waves can become unstable and ultimately break or saturate, and the carried energy momentum gets transferred to higher altitudes. For that reason, gravity waves play a crucial role in the thermal and dynamical structure of the entire Martian atmosphere system.[2] Because the resolution of Martian global atmospheric models is not sufficient to capture all gravity waves, mathematical parameterizations are used to account for their missing effects.

Water vapor is another important component of Martian meteorology. The climate in the lower atmosphere is greatly influenced by the amount and global distribution of water vapor. Water is a complex, radiatively active molecule that is essential to the formation of life, especially in its liquid state, and it could indicate the presence of climatological conditions like Earth's. Although several astronomers in the 19th century claimed to have discovered water in the atmosphere of Mars, the first definitive evidence of water vapor on Mars came in 1963.[10] Despite its currently dry state, Mars has plenty of water, but it is mainly in the vapor and ice states, with the latter located in the polar caps and subsurface ice. Martian water-vapor concentration varies seasonally but hovers around 150 ppm.

Clouds and dust

Water is the main constituent of clouds, which are a vital component of terrestrial weather and life. Clouds consist of micron-scale droplets or ice crystals that are suspended in air and are numerous enough to be observed with the naked eye. They are formed by condensation of water vapor (or other gases) on nucleating hygroscopic particles, which can absorb water. On Earth, clouds are important in the hydrological cycle because they redistribute water to the surface in the form of precipitation. They are radiatively active, absorbing and reflecting high-frequency solar energy at various wavelengths.

Martian clouds are observable by orbiters, rovers, and telescopes (see figures 3 and 4). Satellites can detect them by measuring the decrease in incident radiation caused by atmospheric scattering and absorption or by measuring IR emission at Mars's limb. Rovers and landers equipped with cameras can directly image overlying clouds; the *Mars Pathfinder* lander first imaged them in 1997. They vary in shape, resemble terrestrial cirrus or stratus clouds, and serve as tracers of atmospheric motion and variability. Water and CO_2 ice clouds, which can be spectrally distinguished from each other, are prominent features of the Martian atmosphere.

Mars has a CO_2 cycle that consists of a seasonal global exchange of the gas between the troposphere and the polar ice caps. Although tropospheric temperatures are too warm for CO_2 ice clouds to form, mesospheric temperatures are cold

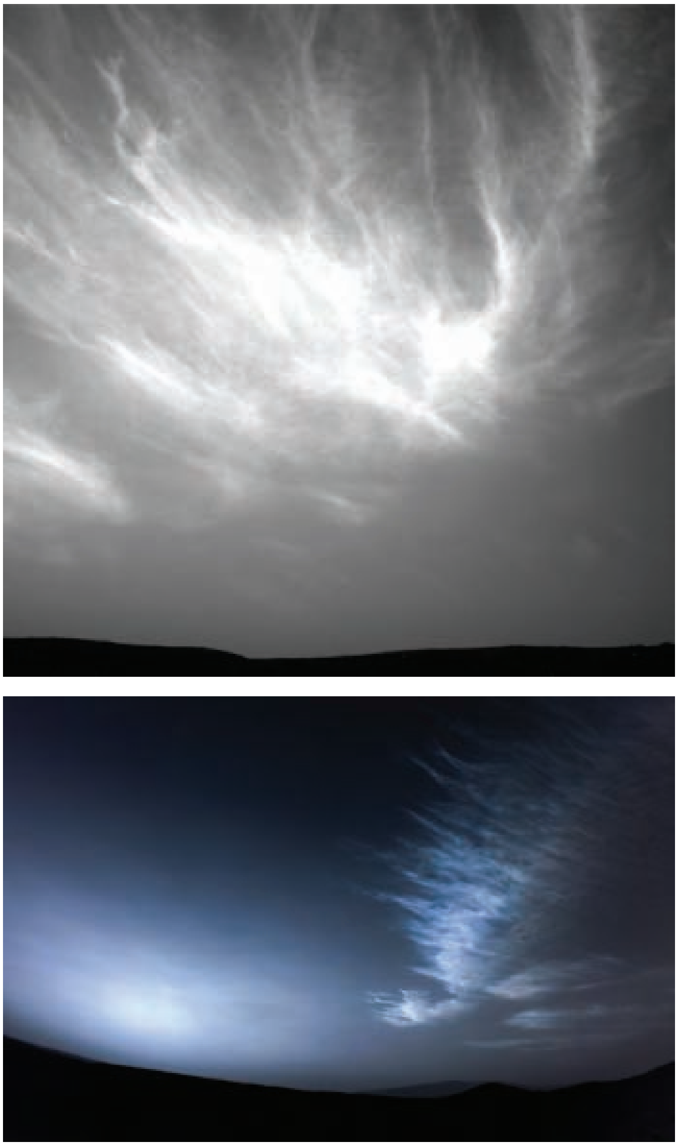


FIGURE 4. IMAGES TAKEN by the *Curiosity* (top) and *Perseverance* (bottom) rovers in 2019 and 2023, respectively, capture cirrus-like cloud structures. (Top image courtesy of NASA/JPL-Caltech; bottom image, NASA/JPL-Caltech/Kevin M. Gill.)

enough to allow for their emergence, especially at night.[11] In the low-pressure environment of the Martian mesosphere, lower temperatures and a higher degree of supersaturation are required for ice crystals to form. By acting as ice-nucleating particles, aerosols provide favorable conditions for the formation of cloud particles.

Martian high-altitude mesospheric clouds are somewhat similar to terrestrial high-latitude polar mesospheric clouds—also known as noctilucent clouds—in their formation altitude and wavy structure. But they also differ in significant ways: Martian clouds have larger ice-crystal sizes, occur across the entire planet, and are distributed over a wider range of altitudes. On Mars, dust storms often raise dust upward, which significantly increases the distribution of aerosols at high altitudes, and thus the storms can be an important factor in high-altitude ice-cloud formation and variability. Gravity waves and solar tides modulate the formation and distribution of mesospheric CO_2 ice clouds by perturbing the background temperature and density. That process can lead to a decrease in local temperature below the CO_2 frost point.

Mars is a dusty planet, and dust is an important atmospheric aerosol that possesses significant radiative properties. Dust can be easily raised and distributed globally, particularly during the perihelion season, when Mars's southern hemisphere experiences summer and the atmospheric circulation intensifies. Martian dust clouds are essentially thunderstorm-like and can rapidly spread across the entire planet. Regional dust storms lasting for a few weeks occur every year; more powerful, long-lasting global major dust storms are rarer.

Understanding dust-storm evolution is essential to landing and communicating with spacecraft on the planet because dust accompanied by fast winds can significantly obstruct sensors and even lead to a temporary shutdown of orbiter and rover operations. In 1971, for example, *Mariner 9* entered Martian orbit only to find the planet's surface completely obscured by a planetwide dust storm. The *Opportunity* rover was permanently shut down in June 2018 by a global dust storm that lowered incoming solar radiation to its solar panels below survivable levels. (The opening image on pages 42 and 43 depicts the beginnings of that dust storm.)

The absorption of solar energy by dust can significantly affect the energy balance of the atmosphere and amplify processes such as deep convection, which involves thermally driven turbulent mixing of the atmosphere from the surface to higher altitudes. The 2018 dust storm increased deep convective activity. It also transported water upward in the atmosphere, thus substantially increasing the water content in the middle atmosphere and causing a rise in the hygropause, the altitude where water vapor condenses into the liquid state and forms clouds.[12]

Space weather and escape

Although Mars has strong localized crustal magnetic fields distributed primarily in the southern hemisphere, it lacks a global magnetic field.[13] For that reason, even though Mars is farther away from the Sun than Earth is, the effects of the solar wind and other space-weather phenomena are more pronounced. Our reference point for understanding space weather is Earth, where the solar wind's interaction with the intrinsic magnetic field and increased particle precipitation into the atmosphere during magnetically active conditions lead to magnetic storms. During those storms, ionization rates increase, and the upper atmosphere is significantly heated because of the increased absorption of solar energy and ion–neutral frictional heating. Strong magnetic fields can confine charged-particle motion and inhibit the direct penetration of charged particles into lower-lying layers.

Other types of space weather, including solar flares and coronal mass ejections, produce intense radiation that affects Mars's atmosphere and ionosphere. In September 2017, one of the strongest—or X-class—flares tripled the ionization radiation within a few tens of minutes, changed the ion density and composition, and increased the ionospheric plasma content.[14] During intense solar-heating events, the atmosphere responds by expanding and redistributing the absorbed heat.

Because atmospheric loss to space may have led to Mars's currently dry and lifeless conditions, the effect of space weather on atmospheric escape is a central subject of research. The degree of escape is typically quantified by an escape flux. Essentially, particles with sufficient translational energy—that is, energy associated with uniform irrotational motion—can escape to space. Around the exobase, particle behavior can be approximated by a Maxwell–Boltzmann distribution. Particles in the high-velocity tail of the distribution that have outward-directed velocities exceeding the escape velocity are likely to be lost to space. Both thermal escape—also known as Jeans escape—and nonthermal escape mechanisms cause atmospheric loss on Mars.

Atmospheric escape can be influenced from above and below. Space-weather effects caused by solar flares, coronal mass ejections, and solar wind–ionosphere interactions can enhance atmospheric emissions, increase ionization and atmospheric temperature, and ultimately lead to a substantial increase in atomic hydrogen escape.[15] In the lower atmosphere, dust-storm-induced thermal changes and gravity-wave upward propagation can lead to enhanced thermal escape around the exobase.[16]

The loss of atomic hydrogen from Mars's upper atmosphere affects the planet's water budget because water photolyzes in the thermosphere and produces hydrogen radicals. Because life on Earth started in water, that loss is relevant to the search for life on Mars, which focuses on water reservoirs under its surface and water vapor in its atmosphere. Mars is thought to have possessed sufficient water on its surface billions of years ago to provide more habitable conditions. But a chain of dynamical and radiative processes, including the thermal loss of hydrogen, appears to have led to the depletion of water reservoirs on Mars.

Living on Mars

Will humanity colonize Mars soon? A deeper understanding of the relationship between Martian meteorology, atmospheric waves, and space weather is a crucial step in planning a potential settlement on the red planet. But many other problems make future human exploration and settlements of Mars an extremely challenging and dangerous endeavor.

To begin with, we would need to get there safely, which would involve surviving nine months of space travel with limited room and resources in a rather unhealthy environment for the human body and mind. Astronauts would be dealing with significant technological, physiological, and psychological challenges. Such a long-range trip has not yet been attempted by humans, although astronauts routinely live for several months in the International Space Station in low Earth orbit.

As an alternative to directly flying to Mars, a base camp could be built on the Moon to serve as a training station for missions to Mars. That is what NASA's Artemis program is attempting (see the article by Michael J. Neufeld, PHYSICS TODAY, December 2023, page 40). Atmospheric entry on Mars has been performed numerous times, but given the highly unpredictive and variable nature of the Martian atmosphere, having humans on board will complicate the procedure.

Mars's surface is hostile to human life. Because the red planet's atmosphere lacks major chemical absorbers, harmful portions of the solar spectrum can penetrate directly to the surface. Furthermore, Mars's lack of a global magnetic field means that charged particles, such as cosmic rays, can reach the surface far more easily than on Earth. Future human settlers would also be confronted with the planet's harsh atmosphere and weather, including dust storms, high-speed winds, and extremely low temperatures. It is a technological mystery whether human suits and settlements can ever be designed to provide continuous protection from those conditions.

Some have suggested terraforming Mars or transforming it into a more habitable planet by means of global-scale environmental engineering.[17] But our current technology does not allow for geoengineering on the planetary scale. An alternative approach would be to adapt the human genome so that it can survive in Mars's environment. But that rather drastic measure would require overcoming biochemical challenges; moreover, it is a subject of intense ethical discussion. A possible compromise might involve future human settlers living permanently in habitats that are technologically equipped to protect them from the Martian environment.[18] In the end, human exploration of the Martian surface will require a major paradigm shift in our perception of humanity's place in the solar system and the universe.

REFERENCES

1. A. S. Medvedev et al., *J. Geophys. Res. Planets* **118**, 2234 (2013).
2. E. Yiğit, *Nat. Geosci.* **16**, 123 (2023).
3. M. H. Fu et al., *Earth Planet. Phys.* **4**, 4 (2020).
4. M. Benna et al., *Geophys. Res. Lett.* **42**, 8958 (2015).
5. E. Dubinin et al., *J. Geophys. Res. Space Phys.* **124**, 9725 (2019).
6. H. Le et al., *J. Geophys. Res. Planets* **127**, e2021JE007143 (2022).
7. S. W. Bougher et al., *J. Geophys. Res. Planets* **120**, 311 (2015).
8. P. Rogberg et al., *Q. J. R. Meteorol. Soc.* **136**, 1614 (2010).
9. E. Yiğit, A. S. Medvedev, P. Hartogh, *Astrophys. J.* **920**, 69 (2021).
10. H. Spinrad, G. Münch, L. D. Kaplan, *Astrophys. J.* **137**, 1319 (1963).
11. R. T. Clancy, B. J. Sandor, *Geophys. Res. Lett.* **25**, 489 (1998).
12. N. G. Heavens, D. M. Kass, J. H. Shirley, *J. Geophys. Res. Planets* **124**, 2863 (2019).
13. M. H. Acuña et al., *Science* **279**, 1676 (1998).
14. E. M. B. Thiemann et al., *Geophys. Res. Lett.* **45**, 8005 (2018).
15. M. Mayyasi et al., *Geophys. Res. Lett.* **45**, 8844 (2018).
16. E. Yiğit et al., *Geophys. Res. Lett.* **48**, e2020GL092095 (2021).
17. C. P. McKay, O. B. Toon, J. F. Kasting, *Nature* **352**, 489 (1991).
18. M. Balistreri, S. Umbrello, *Nanoethics* **17**, 5 (2023).

PT

NEW PRODUCTS

Focus on cryogenics, vacuum equipment, materials, and semiconductors

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

High-vacuum pumping station

Pfeiffer Vacuum now offers its HiCube Neo turbo pumping station for demanding high and ultrahigh vacuum applications in industry and R&D. With pumping speeds ranging from 80 to 800 l/s, the HiCube Neo can be prepared for a wide range of applications that use an array of backing pumps, such as diaphragm, rotary vane, scroll, and multistage Roots. Applications include analyzing gases, calibrating vacuum gauges, and pumping down cryostats. It can also be used in vacuum furnaces. The plug-and-play HiCube Neo is ready for immediate use, and the intuitively operated 7-inch touch screen, remote control via a Web server, and detailed data recording make it user-friendly and efficient. Its design is open to facilitate maintenance and easy access to the components. And it is versatile, so it can be used as a desktop solution or a stand-alone unit. The backing pump's energy-efficient standby mode not only reduces energy use but also operating costs. ***Pfeiffer Vacuum Inc****, 24 Trafalgar Sq, Nashua, NH 03063, www.pfeiffer-vacuum.com*

Cryogenic sample holder

Quantum Machines now offers its QBoard-II cryogenic sample holder to simplify and enhance quantum experiments. Designed for low-temperature spin-qubit chips, general transport experiments, and superconducting circuits, the QBoard-II improves on the performance and usability of the original QBoard. It is more compact and features new designs for the printed circuit board, interposer, and daughterboard. An added shielding lid protects the daughterboard from radiation. The QBoard-II's plug-and-play modular design enables the easy exchange and reuse of sample chips. Support for transmission up to 8 GHz lets users conduct a wide range of experiments. The chip carrier minimizes RF losses and reduces cross talk. To preserve the integrity of experiments, all materials used in the QBoard-II are nonmagnetic. A versatile set of mounting brackets ensures its full compatibility with popular dilution refrigerators. ***Quantum Machines****, HaMasger St 35, Tel Aviv-Yafo, 6721407, Israel, www.quantum-machines.co*

Magnetic research system

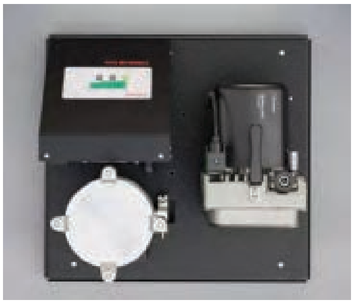
Lake Shore Cryotronics has released the MagRS, a new magnetic research system for material characterization. The base system comes with either a 4- or 7-inch electromagnet base platform and the capability to customize it for specific experimental needs. Options can be specified for vibrating-sample magnetometer and ferromagnetic resonance measurements and for two types of Hall effect measurements for low-mobility material analysis. The AC field Hall option is two times as quiet as the company's 8400 series Hall effect measurement system. The MagRS comes standard with ExactGAP magnet pole gap indexing and GlideLOCK precision sample positioning. Also forthcoming for the MagRS is an electron transport option that integrates Lake Shore's MeasureReady M81-SSM synchronous source measure system for AC and DC field transport measurements. The M81-SSM is easily adaptable for a wide range of characterization applications. ***Lake Shore Cryotronics Inc****, 575 McCorkle Blvd, Westerville, OH 43082, www.lakeshore.com*

Color-changing epoxy

Master Bond's UV15RCL is a low-viscosity, cationic-type UV-curing system with a color-changing feature. Originally red, the material's color becomes clear once exposed to UV light and thus confirms that the light has reached the polymer. That immediate visual feedback offers processing and handling advantages over conventional systems, according to the company. Curing under UV light typically takes 30–60 s with a broad-spectrum UV lamp emitting light with a wavelength range between 320 and 365 nm. The minimum energy required is 20–40 mW/cm^2. The UV15RCL features a high glass transition temperature of 90–95 °C and a service temperature range from −80 to 350 °F. It is not oxygen inhibited and provides light transmission properties and, with a refractive index of 1.517, good optical clarity. It is an excellent electrical insulator with a volume resistivity exceeding 1014 Ω-cm. Suitable for spin coating and for bonding and sealing applications, the UV15RCL adheres well to metals, glass, ceramics, and many plastics, including acrylics and polycarbonates. ***Master Bond Inc****, 154 Hobart St, Hackensack, NJ 07601, www.masterbond.com*

Ultralow-vibration cryostat

Attocube has launched its attoDRY2200 closed-cycle cryostat for ultrasensitive scanning probe microscopy measurements. It can be used in state-of-the-art research in 2D materials, domain walls, multiferroics, superconducting samples with nanostructures, and more. The attoDRY2200 is equipped with a proprietary ultraefficient vibration-damping system, a choice of vector magnets, and automated variable temperature control. It offers a fully automated cooldown from 300 K to a base temperature of 1.65 K without requiring liquid helium. Complemented by an automatic magnetic field control, the superconducting vector magnets make the maximum magnetic field easily accessible over the whole temperature range. Efficient cooling performance via exchange-gas coupling provides initial system cooldown in roughly 15–20 h; the turnaround time during sample exchange is about 5–8 h. The top-loading design enables fast, easy sample exchange and offers a generous sample space of 49.7 mm in diameter. The integrated touch screen allows the desired magnetic field and temperature to be set without using a PC. ***Attocube systems AG**, Eglfinger Weg 2, 85540 Haar, Germany, www.attocube.com*



Environmentally friendly vacuum station

Leybold has improved its TurboLab Core range of entry-level vacuum stations designed for scientific and industrial settings. Combining a primary pump, turbomolecular pump, and controller into one unit, the series especially suits smaller-scale academic research laboratories. The TurboLab Core 90i and 250i can now be supplied with the company's recently introduced ScrollVac 3S as the backing pump. Along with a pumping speed increase from 1.4 to 3 m^3/h, the ScrollVac 3S delivers better ultimate pressure of 0.1 mbar in the same footprint. That and the oil-free, scroll-vacuum technology that the ScrollVac 3S employs demonstrate the company's commitment to environmentally friendly, clean, and carbon-neutral products. With fully replaceable parts and an integrated inlet valve to provide extra protection in the event of a power failure, the new variant is long lasting and reliable and requires fewer maintenance intervals. ***Leybold GmbH**, Bonner Str 498, 50968 Cologne, Germany, www.leybold.com*

Spectroscopic ellipsometer for 4- to 8-inch wafers

With the release of its μSE-2300 microspot spectroscopic ellipsometer, Semilab expands its SE series for metrology in semiconductor device and wafer manufacturing. Featuring a compact, stable platform arrangement with a newly designed metrology head, the μSE-2300 provides high accuracy and improved throughput. It can be combined with a spectroscopic reflectometer and laser ellipsometry in the same system. Various extensions, such as bow and warp metrology, global stress calculation, and model-based dimension (MBD) options, are available. The highly configurable platform can be adapted to support several sample types, including wrapped, thick, or transparent wafers. The versatile μSE-2300 is suitable for various applications, such as thin-film dielectric and semiconductor layer stacks on polished surface substrates. Additional industrial applications involve pattern-capable spectroscopic ellipsometry on high-performance silicon CMOS and III–V devices after deposition and etching processes organic LED displays, and More-than-Moore configurations. ***Semilab USA LLC**, 12415 Telecom Dr, Tampa, FL 33637, https://semilab.com*

High-precision metrology in the mid-IR

The Mid-IR Comb, the latest development in Menlo Systems' family of optical frequency combs, extends precision metrology into the 3–14 μm mid-IR spectral range. That molecular fingerprint signature region allows for unambiguous characterization and provides insights into the chemical composition and reaction kinetics of organic and inorganic materials. The Mid-IR Comb enables highly sensitive and accurate measurements, such as with Fourier-transform mid-IR spectroscopy, nano-FTIR spectroscopy, mid-IR dual-comb spectroscopy, and frequency-locking of mid-IR quantum cascade lasers. The turn-key, all-in-one system is based on Menlo Systems' figure-9 fiber mode-locked oscillator technology and difference frequency generation. The resulting optical frequency comb, free of carrier–envelope offset (CEO), can cover the spectral ranges of 3–5, 5–8, and 8–14 μm and provides up to 200 mW of average optical power within a large spectral bandwidth of 50–300 cm^{-1}. The Mid-IR Comb can also be fully CEO-stabilized in order to obtain comb lines that feature hertz-level linewidths in the mid-IR. ***Menlo Systems Inc**, 6205 Lookout Rd, Unit D, Boulder, CO 80301, www.menlosystems.com* PT

LOOKING FOR A JOB?

Job ads are now located throughout the magazine, alongside the editorial content you engage with each month. Also find hundreds of jobs online at physicstoday.org/jobs

LOOKING TO HIRE?

Enjoy the power of print plus online bundles any time as well as impactful exposure packages & discounts for our special Careers issue each October. Post online-only jobs anytime at physicstoday.org/jobs

Questions? Email us at ptjobs@aip.org

PHYSICS TODAY | JOBS

QUICK STUDY

Mark Boslough is a research associate professor of Earth and planetary sciences at the University of New Mexico and a physicist at Los Alamos National Laboratory, where he models airbursts and their contribution to the impact risk of near-Earth objects.

The threat from cosmic flotsam

Mark Boslough

A power-law distribution of asteroid impacts on Earth spans 13 orders of magnitude in energy. The risk is dominated by low-probability but high-consequence events.

Since its birth 4.5 billion years ago, our home planet has been flying through a sea of debris. The Sun formed when the dense inner part of a nebular disk collapsed and lit itself up by nuclear fusion. Over the following millions of years, most of the solid pieces remaining were pulled together by gravity and conglomerated into planets. As the growing worlds gobbled up more mass, their gravitational influence increased, until most of the original source material was depleted. The leftover bits of solid matter eventually became what we know as comets and asteroids.

The solar system evolved into a reasonably stable configuration, but it never fully settled down. Some objects were big enough to trap their own radioactive heat, and they developed planet-like structures with metallic cores, mantles, and crusts. Eventually, some of those planetesimals slammed into one another and broke up, and the impacts created all sizes of asteroids with different compositions. Over the eons, asteroids continued evolving, with their orbits ever changing through near encounters with planets. Like Earth-bound geology, the evolution has been a slow, never-ending process.

Gravitational encounters, impacts, and other forces have led to a distribution of orbits. The main asteroid belt lies between Mars and Jupiter, but not all asteroids reside there. Many have ended up in the inner solar system and have orbits that cross or pass close to Earth's. They are called near-Earth objects (NEOs), and depending on their paths and locations, some may pose an impact risk.

Cosmic violence shapes the distribution

The ongoing grinding, smashing, and breaking of asteroids have also resulted in an emergent phenomenon: a power-law size distribution. Clark Chapman and David Morrison, in their 1989 book *Cosmic Catastrophes*, described the nonuniformity of asteroid sizes as an inherent property of the 3D universe. The chances of collisional fragmentation of a given asteroid depend on how large it is. When a big asteroid breaks up, the total surface area of material increases, and the more numerous fragments have a higher total chance of being involved in other collisions, which would cause them to break up into even smaller pieces.

Figure 1 shows the size distribution of near-Earth asteroids. The biggest ones, shown on the right-hand side of the graph, collide with Earth every few hundred million years on average. A smaller one, about 10 km in diameter, formed the Chicxulub impact crater when it wiped out the dinosaurs 66 million years ago and changed our planet forever. The probability of a similar impact catastrophe in a given year is about one in a hundred million, the reciprocal of the mean impact interval. That's greater than the odds of winning a Powerball jackpot.

At the other end of the spectrum are objects a few meters in diameter, which US government sensors observe exploding in the atmosphere as fireballs, also called bolides, several times every year. They are frequent but inconsequential in terms of risk—the most likely victim being a person, car, or house.

On the continuum of asteroid size, a threshold exists in which the explosions are as big as nuclear detonations and close enough to Earth's surface to kill people and destroy infrastructure. They are low-altitude airbursts, like the simulated example shown in figure 2, and are rare on the time scale of human lives. The meteor that fell over Chelyabinsk Oblast, Russia, in 2013 was about 20 m in diameter. It released about

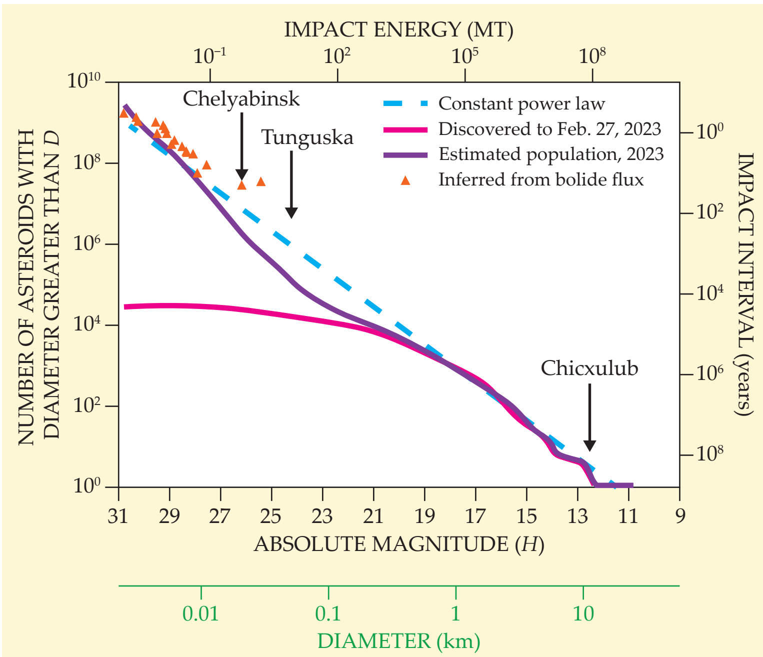


FIGURE 1. NEAR-EARTH ASTEROID size–frequency distribution is shown in units of asteroid diameter *D* and absolute magnitude *H*, a measure of intrinsic brightness. Smaller asteroids are dimmer and the hardest to discover, so their discovered fraction gets smaller as the size decreases, resulting in a flattening of the magenta curve. (Courtesy of Alan Harris.)

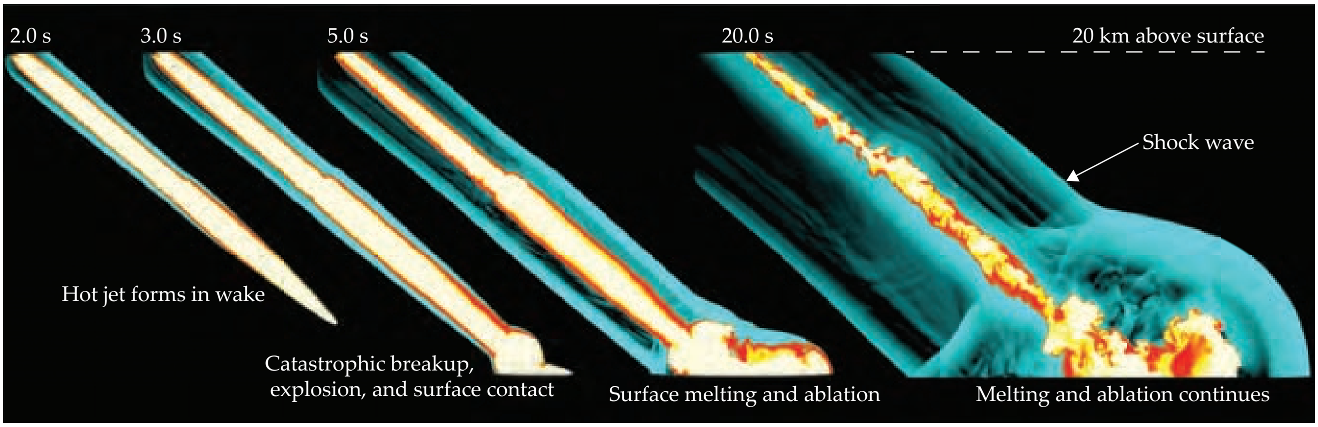


FIGURE 2. A SIMULATION. A 100 m asteroid enters Earth's atmosphere at a 35° elevation angle and traveling 14.2 km/s. It has a 30 megaton kinetic yield and is shown passing beyond a point 20 km above Earth. In my numerical simulations, each frame is one in a sequence, a few seconds apart. (The time in each one is the number of seconds since the asteroid passed the 20 km point.) In the second frame, the airburst occurs 2 km above ground. The resulting vapor jet keeps descending in subsequent frames until it reaches the surface.

a half megaton of explosive energy and injured more than a thousand people (see my article with David Kring, Physics Today, September 2014, page 32). Airbursts of that size happen once every 50 years on average.

The 1908 explosion in Russia known as the Tunguska event is probably an order of magnitude more energetic (and likely several megatons), but that estimate is uncertain because the event happened at a time and location for which observational and instrumental data are sparse. The best data come from the physical evidence left on the ground, where trees were blown down over an area spanning more than 2000 square km—the size of a large metropolitan area. We can expect events of that size to happen with a mean interval of about 500 years.

Distribution shapes the risk

The first attempt at a quantitative probabilistic NEO risk assessment was published in 1994 by Chapman and Morrison. Their estimates of airburst damage were based on nuclear weapons' effects and scaling laws. But that method breaks down for larger asteroids because such global effects as climate change come into play. Chapman and Morrison extrapolated their estimates up to a global-catastrophe-threshold asteroid size between 0.5 and 3 km in diameter, above which they assumed a quarter of the world's population would die. That estimate provided a crude framework for deciding how to begin reducing the danger.

By integrating their risk curve, Chapman and Morrison estimated a few thousand fatalities per year. That result is counterintuitive because there is no direct evidence that anyone has ever been killed by an asteroid. The risk is dominated by low-probability, high-consequence events. Fatal impacts are rare but expected to kill many people when they happen.

The best way to reduce the risk is to prevent large impacts. Preventing a catastrophic impact requires finding all the NEOs above the global catastrophe threshold, so a survey program was established by a 1998 NASA directive to discover 90% of NEOs greater than 1 km in diameter. Fortunately, there are only about 1000 NEOs of that size. Because they are the biggest and brightest in the sky, they are also the easiest to discover.

Eliminating catastrophic risk with surveys uses the same principle as looking both ways before crossing the street. Situational awareness doesn't by itself change the probability of impact. An NEO will either collide with Earth on some specified time interval or it won't. Observation creates the opportunity to take preventive action to mitigate the risk if something is discovered to be on a collision course.

The method of choice for planetary defense is to deflect an asteroid from its collision course by sending a spacecraft to collide with it and changing its velocity. That option is available only if the asteroid is discovered well in advance, because there must be sufficient time for the asteroid to drift away from where it would otherwise be at the time it crosses Earth's path.

Astronomical surveys may have reduced our assessment of the likelihood of a global- or continental-scale catastrophe by an order of magnitude, to an estimate of about 100 fatalities per year. Recent advances in hydrocode models, however, suggest that severe airburst effects, which also accompany crater-forming impacts, may have much more severe consequences than we had imagined and that the global catastrophe threshold might be triggered by smaller asteroids than we had thought. So the probability-based reduction in catastrophic risk may have been partially offset by an understanding-based lowering of the catastrophe threshold.

As the risk from larger asteroids shrinks, the relative hazard shifts to small impacts or airbursts that are much more frequent and difficult to avoid. Barring the discovery of a large NEO on a collision course, mitigation methods will shift from deflecting the asteroid to evacuating people from the impact zone, an activity that the Federal Emergency Management Agency and NASA are practicing with exercises involving simulated strikes. A new space-based IR telescope, the *Near-Earth Object Surveyor*, will be an essential tool for reducing the remaining risk and providing early warning.

Additional resources

►C. R. Chapman, D. Morrison, *Cosmic Catastrophes*, Plenum Press (1989).

►C. R. Chapman, D. Morrison, "Impacts on the Earth by asteroids and comets: Assessing the hazard," *Nature* **367**, 33 (1994).

►M. Boslough, P. Brown, A. Harris, "Updated population and risk assessment for airbursts from near-earth objects (NEOs)," in *2015 IEEE Aerospace Conference*, IEEE (2015).

►M. Boslough et al., "FEMA asteroid impact tabletop exercise simulations," *Procedia Eng.* **103**, 43 (2015).

►A. W. Harris, P. W. Chodas, "The population of near-earth asteroids revisited and updated," *Icarus* **365**, 114452 (2021). PT

BACK SCATTER

Viewing the brain's nanostructure

To understand how the brain's structure affects various functions, you need a detailed map of the neurons and their connections. Viren Jain of Google Research, Jeff Lichtman of Harvard University, and colleagues made such a map by using volume electron microscopy and advanced computational techniques. The approach's rapid imaging and automation capabilities made it possible to reconstruct a 1 mm^3 volume of tissue from a human brain's outer layer. The sample had been surgically removed from a patient so that neurologists could reach an underlying region that is a source of epileptic seizures. The volume's nanoscale resolution is fine enough to make visible individual cellular elements and synapses—the communication sites between neurons—and the encompassed area is large enough to include neural circuits made of tens of thousands of neurons and millions of synapses.

This image shows thousands of neurons whose cell bodies are shaped loosely like pyramids. The neurons range in length from 10 μm (blue) to 30 μm (red). Neurons receive electric signals via dendrites—the thick branches extending from the neurons—and send signals using axons, the longer, thinner fibers between the cells. Researchers suspect that autism, schizophrenia, and other developmental disorders may alter the structure of neural circuits. More analyses of the data may be forthcoming: The researchers have made the 1.4 petabytes of reconstruction data publicly available online. (A. Shapson-Coe et al., *Science* **384**, eadk4858, 2024; image made by Daniel Berger, courtesy of Google Research and the Lichtman Lab at Harvard University.) —AL

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