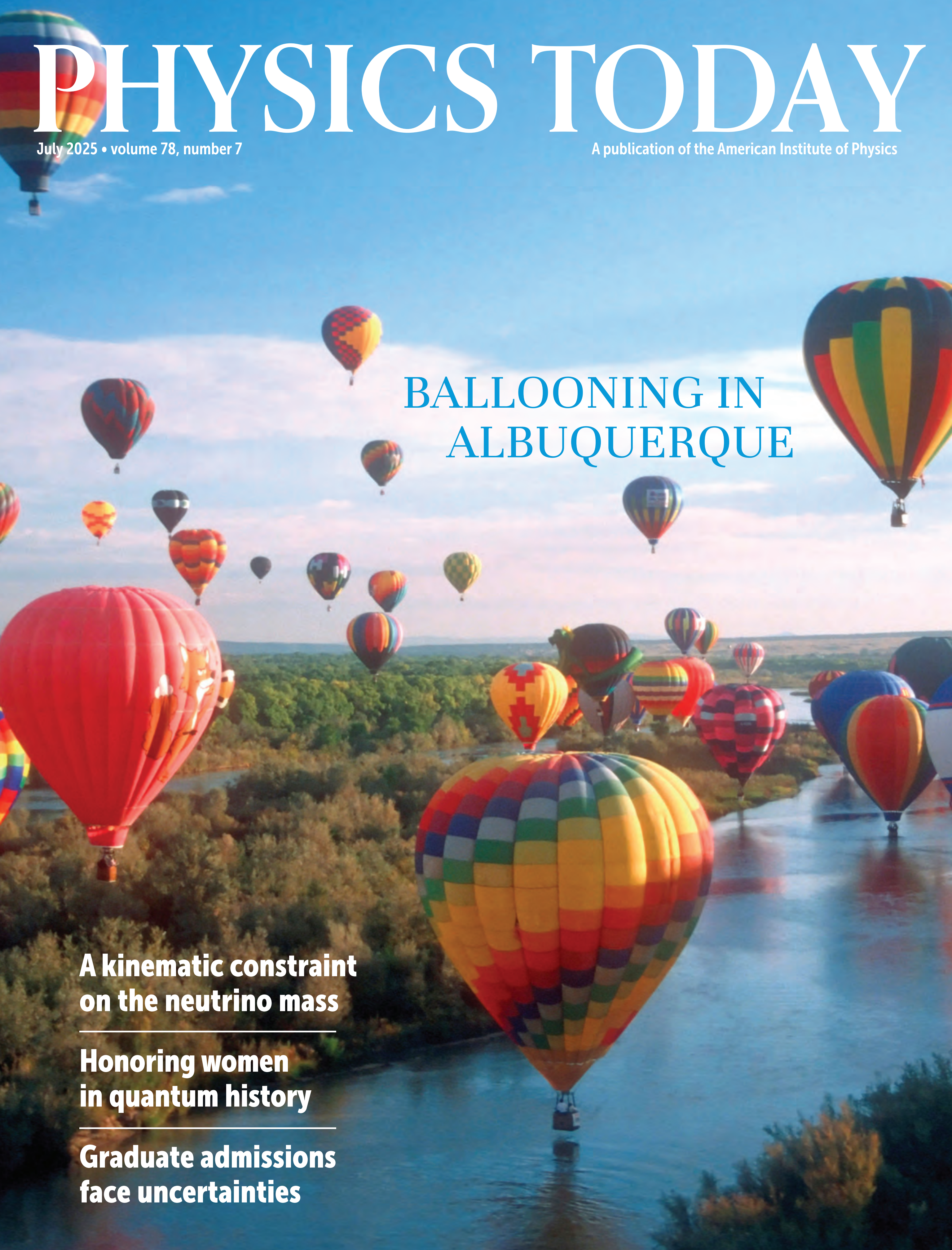


PHYSICS TODAY

July 2025 • volume 78, number 7
A publication of the American Institute of Physics

BALLOONING IN ALBUQUERQUE

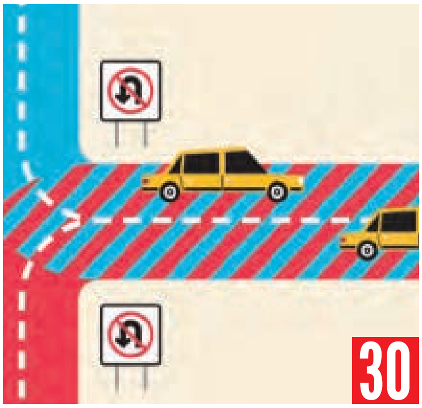
A kinematic constraint on the neutrino mass

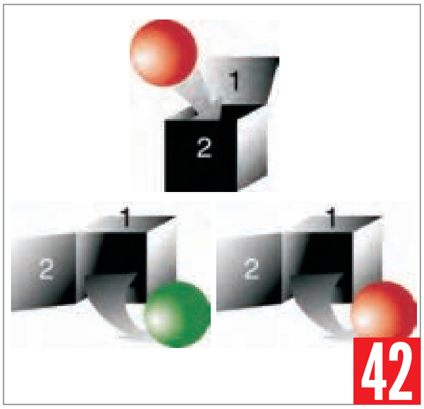
Honoring women in quantum history

Graduate admissions face uncertainties

PHYSICS TODAY

July 2025 | volume 78 number 7





FEATURES

ON THE COVER: Aficionados of hot-air ballooning flock to Albuquerque, New Mexico, for an experience that's not possible in most other parts of the world. From a launch site to the north of the city, balloon pilots fly south across the metropolitan area. Then, astonishingly, they turn around and return to the starting point. To learn about the geographic and atmospheric conditions that enable such a feat, turn to the Quick Study on **page 54**. (Image by Mark Newman/Rainbow/RGB Ventures/SuperStock/Alamy Stock Photo.)

Recently on PHYSICS TODAY ONLINE

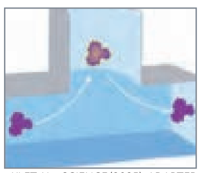
www.physicstoday.org

CERN

Europe's next collider?

The Future Circular Collider (illustrated on the map above) is among the options that Europe's particle-physics community is considering as the successor to CERN's Large Hadron Collider. At stake in the pending decision are tens of billions of dollars and decades of construction, development, and research.

physicstoday.org/Jul2025a

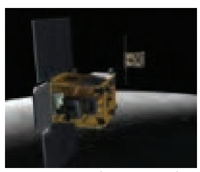


XI ET AL., *SCIENCE* (2025), ADAPTED

Probing biomolecules

A new technique allows researchers to swiftly and easily glean basic structural information about biomolecules that are put in solution to mimic their natural environments. Structural properties are gathered through measurements of how quickly a molecule escapes from a trap etched into a microfluidic channel.

physicstoday.org/Jul2025b



NASA/JPL-CALTECH/MIT

The Moon's mantle

An analysis of data from a 2012 lunar gravity-mapping mission suggests that the Moon's mantle is warmer on one side than the other. The asymmetry, which is likely caused by the distribution of radioactive elements, could explain some of the defining surface features on the Moon's near and far sides.

physicstoday.org/Jul2025c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 78, 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.

Alliance for Audited Media

PHYSICS TODAY

www.physicstoday.org

DEPARTMENTS

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

Editor-in-chief
Richard J. Fitzgerald *rjf@aip.org*

Managing editors
Andrew Grant *agrant@aip.org*
Johanna L. Miller *jlm@aip.org*

Art and production
Freddie A. Pagani, art director
Nathan Cromer
Abigail Malate
Three Ring Studio

Editors
Ryan Dahn *rdahn@aip.org*
Jenessa Duncombe *jduncombe@aip.org*
Laura Fattaruso *lfattaruso@aip.org*
Toni Feder *tf@aip.org*
Abby Hunt *ahunt@aip.org*
Alex Lopatka *alopatka@aip.org*
Gayle G. Parraway *ggp@aip.org*

Assistant editor
Nashiah Ahmad *nahmad@aip.org*

Digital operations
Greg Stasiewicz *gls@aip.org*

Editorial assistant
Tonya Gary

Contributing editors
Mitch Ambrose
Hannah Daniel
Bob Holmes
Andreas Mandelis
Lindsay McKenzie
Jacob Taylor
Clare Zhang

Sales and marketing
Christina Unger Ramos, director *cunger@aip.org*

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

AIP

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

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

FROM THE EDITOR

A new Physics Today is coming

Richard J. Fitzgerald

As so many current events are showing, the scientific world can evolve quickly, much faster than a monthly publication can keep up with. That is certainly a significant motivator, but not the only one, for major changes that Physics Today will be making later this year.

PT was established more than 75 years ago to counter the accelerating growth and fractionation that the field of physics was experiencing following World War II. (For more on those early years, see the article by David Kaiser, Physics Today, May 2018, page 32, commemorating *PT*'s 70th anniversary.) Increasing research specialization and compartmentalization risked the erosion of the field's sense of community. By providing news and feature articles across the gamut of physics-related topics, *PT* was intended to give readers an appreciation of what was going on outside their specialty and thus establish and maintain a sense of common identity.

Fostering that sense of community still underpins *PT*'s mission today: to be a unifying influence on the physical sciences by cultivating a shared understanding, appreciation, and sense of belonging among physical scientists. We strive to acknowledge our commonalities, honor our history, and celebrate our differences.

In a world in which news, progress, and change are spreading faster than ever, having a trusted, reliable source of accurate, curated information that matters is even more essential. And that information will soon have a new home: a *PT* website that will provide a modern, magazine-like digital experience. No matter how you reach us—or we reach you—we will, we will make it easy to discover recent research breakthroughs and applications, indulge your curiosity, find inspiration from Q&As, and locate resources to advance your career.

At our core, we will continue to cover twin aspects of the physical sciences enterprise: the richness of the science and the shared experience of being a scientist. Amid today's unprecedented upheaval in the sciences and science education, that shared experience is taking on extra urgency and importance. Our new online home will better allow you to find the information that's relevant to you—whether it's *PT's* own content, reports on science policy and statistics from other parts of the American Institute of Physics (publisher of Physics Today), or material from the Member Societies that make up the AIP Federation.

PT's news stories and feature articles will first appear on our website, and most will then appear in the next issue of the monthly magazine. That way, you won't have to wait a month for the content you want.

Just as importantly, the platform will make it easy to engage both with our content and with others in the community. For the majority of *PT*'s existence, letters to the editor provided an important conduit for readers to share their thoughts. These days, however, digital channels reaching a receptive audience are essential. Indeed, the exchange of ideas remains central to building a strong, shared culture and a welcoming community.

As we roll out these changes—and additional ones in the future—we will be seeking the input of you, our readers, to understand how we can best serve you and meet your needs. We are currently recruiting volunteers for focus groups; if you're interested in sharing your perspectives, please visit https://www.aip.org/pt-focus-group-sign-up. And if you receive a link later this year to an online reader survey, please take it. But as always, you don't need to wait to let us know your thoughts: You can reach us any time at https://contact.physicstoday.org or pteditors@aip.org. PT

READERS' FORUM

Comments on early space controversies

David Cummings and Louis Lanzerotti's feature "Early debates in space science" (Physics Today, February 2025, page 38) provides fascinating accounts of observations and proposals pertaining to the solar wind and the heliosphere from the 1950s onward. I was disappointed that the study of auroras was not included in those accounts because our understanding of them is also connected to the solar wind.

Kristian Birkeland, between 1895 and 1916, published many accounts of his theories of the aurora in scientific papers and books. His concept was that they are generated by charged particles emitted by the Sun and captured by Earth's magnetic fields. He commented that the particles (electrons) and ions coming from the Sun that are not captured are swept away from Earth and continue through the solar system, and he elaborated on this concept in the 1913 edition of his book *The Norwegian Aurora Polaris Expedition 1902–1903*. He thus connected the aurora with what we now call the solar and stellar winds, although he did not use those modern terms.

Birkeland's work generated much debate and was vigorously opposed by geophysicist Sydney Chapman. A very readable account of Birkeland, his work, the controversy it generated, and his tragic life is available in the 2001 book *The Northern Lights* by Lucy Jago.

Bruce McKellar
(bhjmckellar@mac.com)
University of Melbourne
Melbourne, Australia

~~~

David Cummings and Louis Lanzerotti's article "Early debates in space science" (Physics Today, February 2025, page 38) tells the stories of five early questions in astrophysics. One of those big questions was about the nature of gamma-ray bursts: whether they come from within the Milky Way or beyond it. As the authors discuss, a great debate on the subject was held in 1995. The debate papers were contained in a special issue[1] of *Publications of the Astronomical Society of the Pacific*, for which I was managing editor.

The resolution of the issue came two years later, but not quite in the way that Cummings and Lanzerotti describe. A

THE AURORA BOREALIS in 2013. (Photo from LCDR Gary Boone, NOAA.)
~~~

gamma-ray burst, named GRB 970508, occurred on 8 May 1997 and was detected by the *BeppoSAX* satellite, which provided a fairly accurate celestial position. I used a 0.9 m telescope at Kitt Peak National Observatory to image the location on two successive nights, resulting in the detection of a faint optically variable source within the error box.

Following my announcement, which gave accurate coordinates of the object,[2] Charles Steidel of Caltech was able to obtain its spectrum at the W. M. Keck Observatory.[3] He reported that the afterglow has a redshift z of 0.835 and settled once and for all that GRBs indeed lie at cosmological distances.

As Cummings and Lanzerotti's article recounts, Bohdan Paczyński had been the advocate for cosmological distance at the great debate. When I emailed him in the early morning to inform him of the results and to congratulate him on being right, he told me that he believed that he would allow himself a drink that evening.

References

1. Special issue, "The 75th anniversary astronomical debate on the distance scale to gamma-ray bursts," *Publ. Astron. Soc. Pac.* **107** (1995).
2. International Astronomical Union Circular No. 6654, 10 May 1997.
3. International Astronomical Union Circular No. 6655, 11 May 1997.

Howard E. Bond
(heb11@psu.edu)
Pennsylvania State University
University Park

~~~

When reading David Cummings and Louis Lanzerotti's article on "Early debates in space science" (Physics Today, February 2025, page 38), I was surprised to see their account of the solution to the mystery of where gamma-ray bursts come from.

Cummings and Lanzerotti mention the 1995 debate, held at the Smithsonian Institution's National Museum of Natural History, on whether gamma-ray bursts are galactic or extragalactic. They state that the debate "did not resolve the dispute" but rather a "combination of space- and ground-based observations two years later did." The authors mention Jan van Paradijs and his students, who in 1997 "were able to associate a gamma-ray burst with a specific galaxy" but unable to measure its emission-line spectra. They also mention a group led by Mark Metzger, who found a gamma-ray burst that occurred simultaneously with an optical flash, and resulting measurements "established beyond doubt that the burst sources were outside our galaxy."

The article does not mention that it was the Italian–Dutch satellite *BeppoSAX* that detected and promptly, and accurately, localized a gamma-ray burst that occurred on 28 February 1997. In addition, with the same satellite, it was possible to discover the first x-ray counterpart of a gamma-ray burst event.[1] The *BeppoSAX* team, of which I was one of the leaders, rapidly distributed the event coordinates in the International Astronomical Union circular. That made it possible for Jan van Paradijs and colleagues to discover an optical transient that had a position consistent with the gamma-ray burst x-ray counterpart.[2]

Also, the determination of the first gamma-ray-burst redshift by Metzger's group[3] was the result of the *BeppoSAX*'s detection and prompt, accurate localization of another event, GRB 970508. With the same satellite, it was also possible to discover its x-ray counterpart (that is, its x-ray afterglow).[4] And thanks to the prompt alert of our collaborators in Caltech, led by Shri Kulkarni, and those at the Very Large Array radio telescope, led by Dale Frail, it was possible to discover the optical and radio counterparts and to measure its redshift.[3, 5]

For a more extended history of these discoveries, see my recently published review in reference 6.

References

1. E. Costa et al., *Nature* **387**, 783 (1997).
2. J. van Paradijs et al., *Nature* **386**, 686 (1997).
3. M. R. Metzger et al., *Nature* **387**, 878 (1997).
4. L. Piro et al., *Astron. Astrophys.* **331**, L41 (1998).
5. D. A. Frail et al., *Nature* **389**, 261 (1997).
6. F. Frontera, *Universe* **10**, 260 (2024).

Filippo Frontera
(filippo.frontera@unife.it)
University of Ferrara
Ferrara, Italy

► **Cummings and Lanzerotti reply:** We thank Bruce McKellar, Filippo Frontera, and Howard Bond for their comments in response to our article on early debates in space science that appeared in the February 2025 issue of Physics Today.

As McKellar states, there was certainly a great controversy exercised by Sydney Chapman over the existence of Kristian Birkeland's geomagnetic field–aligned currents. That important space science controversy was resolved by measurements taken by the *1963 38C* satellite[1] and analyzed by one of us (Cummings) with Alexander Dessler.[2] The controversy was not related to the debate, described in our feature, on whether the magnetosphere is open or closed. Alv Egeland and William Burke cover the life and career of Birkeland in more detail in their 2005 book, *Kristian Birkeland: The First Space Scientist*.

We particularly thank Frontera and Bond for adding personal details to the story about the determination of the distance scale of gamma-ray bursts. They cite our failure to mention the role of the *BeppoSax* satellite and its scientific team. Unfortunately, the word limit for our Physics Today article forced us to make difficult choices as to what to include. The contributions of *BeppoSax* and its team and the observation of GRB 970508 and its afterglow in its host galaxy are described in chapter 9 of our 2023 book, *Scientific Debates in Space Science: Discoveries in the Early Space Era*.

References

1. A. J. Zmuda, J. H. Martin, F. T. Heuring, *J. Geophys. Res.* **71**, 5033 (1966).
2. W. D. Cummings, A. J. Dessler, *J. Geophys. Res.* **72**, 1007 (1967).

W. David Cummings
(wcummings@usra.edu)
Universities Space Research Association
Washington, DC

Louis J. Lanzerotti
(louis.j.lanzerotti@njit.edu)
New Jersey Institute of Technology
Newark

CONTACT PHYSICS TODAY

Letters and commentary are encouraged and should be sent by email to **ptletters@aip.org** (using your surname as the subject line), or by standard mail to Letters, Physics Today, American Center for Physics, 1 Physics Ellipse, College Park, MD 20740-3842. Please include your name, work affiliation, mailing address, email address, and daytime phone number on your letter and attachments. You can also contact us online at **https://contact.physicstoday.org**. We reserve the right to edit submissions.
~~~

Eight million years of Arabian climate were not all dry

The Arabian Desert is a barrier to animal and plant migration between Africa and Eurasia, but that hasn't always been the case.

FIGURE 1. A CAVE ENTRANCE disrupts the vast expanse of rock and sand in the Arabian Desert. This and other caves in central Saudi Arabia contain stalagmites and stalactites—rocks formed by the seepage of water through soil and limestone—that reflect humid phases in the peninsula's past. (Photo by Paul Breeze.)

If you've ever been in a cave full of long, spindly stalagmites and stalactites, you might recall a sound that frequently interrupts the otherwise cavernous silence: Drip, drip. The flow of water is the very reason the rock formations are there. When water saturated with calcium carbonate trickles into a cave, carbon dioxide degases from the water, and calcite crystallizes onto the cave's surfaces. Layers slowly build up and, like tree rings, offer a record back through time.

The conditions needed to create those water-deposited cave rocks, collectively known as speleothems, are not what anyone would expect to find in the middle of a vast, barren desert such as the one in figure 1. For water rich in calcium carbonate to arrive at a cave, water must first pick up carbon dioxide by traveling first through soil that's rich with the remains of dead plants. That process makes the water acidic enough to dissolve calcite from limestone. And yet, from caves in the middle of the Arabian Desert, about 200 km north of Riyadh in Saudi Arabia, scientists have retrieved dozens of speleothems, including the ones seen in figure 2, that tell of a time when the peninsula looked much different—and greener—than it is today.

Now Monika Markowska (currently at Northumbria University in the UK but formerly with the Max Planck Institute for Chemistry in Germany) and colleagues have published their initial findings about the climate history recorded in those speleothems.[1] The researchers were stunned when they found that their samples date back nearly 8 million years, roughly an order of magnitude older than what they had expected, and about 7 million years longer than any existing climate record from the peninsula. They identified distinct humid periods, centered around 7, 4–3, 2, and 1 million years ago, during which the speleothems grew.

Back in time

The barrier posed by the vast expanse of the harsh Saharo-Arabian Desert is part

FIGURE 2. SPELEOTHEM SAMPLES from Mossy Cave in the Arabian Desert are collected by Alexander Budsky **(left)** and Hubert Vonhof **(right)**. The cave's name was inspired by mossy vegetation that grows at its mouth due to moist air that escapes from inside the cave. The rocks grew during humid intervals that span the past 8 million years. (Photo courtesy of the Green Arabia project.)

of the reason Africa has so many unique, large animals. Recent research has documented that the desert's presence goes back 11 million years.[2] The Arabian Peninsula is one of just a few geographies, which also include the Nile Valley and the narrow crossing at the Strait of Gibraltar, that archaeologists believe could have once served as a conduit for early human ancestors out of Africa and into Eurasia. Episodes when water-dependent animals took up residence in the area are reflected in the fossil record from the peninsula during two distinct time periods, the past 500 000 years and 7 million years ago, though little has been known about the climate of the intervening time span.

The long record collected from the speleothems can, ironically, be partially attributed to the desert environment: If wetter climate periods were more frequent or lasted longer, the older speleothem samples could have been covered up by younger ones. In the dynamic landscape of the desert, cave entrances, such as the one seen in figure 1, are routinely buried and uncovered by the movement of dunes. That has also likely aided the samples' preservation.

The standard method for dating speleothems relies on the ratios of parent isotope uranium-234 to its decay product thorium-230. It's a handy tool for speleothem dating because uranium dissolves in water but thorium does not. So when a layer of calcite precipitates out of water, it contains uranium but no thorium. The fresh layer is a clean slate from which to measure the passage of time via the accumulation of thorium. But that technique can go back only about 600 000 years. After that, ^{230}Th, a radioactive isotope, reaches an equilibrium value because it starts to decay at the same rate that it accumulates.

When Markowska started to investigate how old the speleothems are, she ran into a problem: 90% of the measurements were out of range for U–Th dating (see figure 3 for an example). "That was a bit of a disappointment at first," says Hubert Vonhof, who leads the isotope geochemistry group at the Max Planck Institute for Chemistry and was part of the research team.

Undeterred, the team turned to another radioactive timekeeper: uranium–lead dating. That approach relies on two radioactive decay chains: ^{238}U to ^{206}Pb and ^{235}U to ^{207}Pb. Age calculations require corrections for the initial ratios of uranium isotopes in a sample. The researchers calculated averages for those initial values using the younger sections of rocks dated with U–Th ratios. The need for such corrections yields larger uncertainties compared with U–Th ages.

Uranium–lead dating can also be challenging, in part because lead is everywhere, so samples are easily contaminated by something as simple as dirt. "It was a little bit lucky that uranium–lead dating worked. The samples were quite pristine," says Vonhof. Contamination from dirt can often be seen in the lab when speleothems are processed, but excess lead also stands out in the isotope measurements. The use of laser ablation, rather than standard solution chemistry, to perform spectroscopy also enabled faster U–Pb dating of more samples.

Fossil water

More luck came in the form of water that was directly sourced from the past; it was trapped in the calcite crystal lattice of the speleothems. The source of water can be fingerprinted by the relative ratios of ^{18}O to ^{16}O and ^{2}H to ^{1}H and compared with global databases. Stable isotopes of hydrogen and oxygen from the water show that it originated from a seasonal, southerly source of moisture brought in by monsoonal winds moving through the Arabian Sea and the Gulf of Aden.

In contrast, the sparse, modern rainfall comes from a combination of northern and southern sources, with no strong seasonal signal. As the source of rainfall at the Arabian cave sites changed slowly over time to more northern sources, the

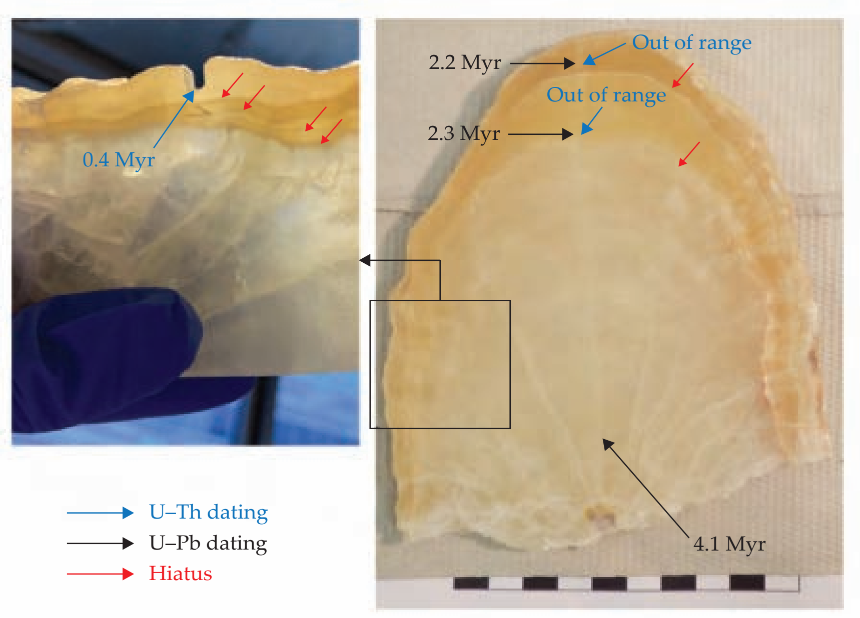


FIGURE 3. A STALAGMITE CROSS SECTION from Mossy Cave is one of many used to extract radioisotope ages that illuminate the timing of humid phases on the Arabian Peninsula. Uranium–thorium ratios (blue) are analyzed from chemical solution and can measure back only about 600 000 years. Uranium–lead ratios (black) measured with laser ablation show that samples date back several million years (Myr). Red arrows point to hiatus periods during which there was no mineral growth on the stalagmite for an extended time. (Photos adapted from ref. 1.)

overall Arabian climate was getting drier. Even the humid periods became less humid. The aridification trend is reflected not only by the oxygen isotope measurements from the trapped water and calcite, because heavier oxygen isotopes are enriched during drier periods, but also by the crystal structure of the calcite: During the older wet phases, larger calcite crystals could grow.

One of the geographically nearest paleoclimate records to date is from an area in the Zagros Mountains in Iran, about 1000 km north of the Arabian cave sites; the sedimentary rocks there show hyperaridity over the period spanning 5.6–3.3 million years ago.[3] The new data that show a humid phase from 4 to 3 million years ago complicate that picture and illuminate some of the geographic variability of the climate at the time.

Miriam Belmaker, a paleoanthropologist at the University of Tulsa in Oklahoma, says that the speleothem data open a new line of consideration for the timing of and paths taken by hominins migrating directly across the southern or central Arabian Peninsula rather than north through the Levant region. (Hominins are members of the Hominini taxonomic tribe that includes humans and all our bipedal ancestors.) "Up to now, the dispersal routes through Arabia were not considered as viable," she says. The Nile has been regarded as the more attractive passageway from Africa into Eurasia, but there's no archaeological or paleontological evidence that correlates the region to the earliest established hominin dispersal about 2 million years ago.

"Humid periods in Arabia, where you can detect enough precipitation to have formed speleothems, could have supported a dispersal of humans," says Belmaker. She points out that the humid period in the Arabian Peninsula from 4 to 3 million years ago could be especially important. There are archaeological sites in Asia, such as one at the base of the Himalayas that goes back 2.5 million years,[4] that contain stone tools and cut bones—evidence for an even older hominin presence outside Africa. But without a direct migration route in mind, those artifacts haven't been widely adopted as significant findings worth further investigation. Belmaker expects the new climate data to motivate deeper exploration. Of course, dispersal and mixing of other animals and plants would also have been facilitated by those humid phases.

Aridification

Ratios of stable carbon isotopes from the speleothems hint at correlations between climate shifts and changing vegetation types in the region, toward vegetation more adapted to dry climates. But the researchers are hesitant to draw strong conclusions from the isotopes; instead, they are investigating further. The speleothems contain some preserved pollen, another rare find for the region. Analyses of the pollen are underway and could provide a stronger picture of the evolving ecosystem and climate.

The age data extracted from the cave rocks don't provide strong constraints on the exact duration of wet and dry phases. It's possible that those phases occurred over even narrower ranges of time: tens to hundreds of thousands of years. Data from hundreds of speleothems would need to be collected to provide robust population statistics that could narrow down the ranges, but analyzing so many samples would be prohibitively expensive. The longest apparent dry period in their data, from 6.3 to 4.1 million years ago, corresponds with a well-known climate event in the broader region known as the Messinian salinity crisis, during which the Mediterranean Sea nearly dried up.

In the portion of the speleothem record spanning the past million years, as the climate of the peninsula became progressively more arid, calcite stopped forming on the speleothems and was replaced by gypsum. A rind of gypsum left behind signifies the end of the standard mechanism for speleothem formation: Less water is needed, and it doesn't have to travel through organic material to become acidic and dissolve limestone; gypsum can be dissolved more easily. Says Markowska, "We can really see the hydrological change reflected in the speleothems through time."

Laura Fattaruso

References

1. M. Markowska et al., *Nature* **640**, 954 (2025).
2. A. J. Crocker et al., *Nat. Geosci.* **15**, 671 (2022).
3. M. Böhme et al., *Commun. Earth Environ.* **2**, 85 (2021).
4. D. Cauche et al., *L'Anthropologie* **125**, 102846 (2021).

Kinematic measurements are closing in on the neutrino mass

The lightest massive particles are clear evidence of physics beyond the standard model. But the masses and their implications remain to be understood.

The standard model of particle physics has an admirable track record. It accurately describes how, beneath the familiar world of protons, neutrons, electrons, and photons, there lurks a hidden zoo of quarks and gluons, leptons and gauge bosons. It even predicted the existence and behavior of previously unknown particles—the top quark and Higgs boson—decades before they were experimentally observed.

But the model is not perfect. It's frustratingly silent on many questions, including why particles' masses and interaction strengths are what they are. It sits in awkward tension with some observations, such as the fact that the universe contains much more matter than antimatter.

And sometimes its predictions are wrong. The model suggests that neutrinos should have zero mass. But we know they do not.

It's reasonable to think that at least some of the standard model's weaknesses may be connected, and that poking at one of the chinks in the model's armor could yield the insight needed to build a more complete and comprehensive theory of the universe. Toward that end, researchers working on the KATRIN (Karlsruhe Tritium Neutrino) experiment in Germany—part of the apparatus of which is shown in figure 1—have been chipping away at the range of possible masses for the electron neutrino by carefully analyzing the energy spectrum of electrons emitted in beta decay.

The latest KATRIN result, based on a quarter of the data that the experiment will ultimately collect, still provides only an upper bound on the neutrino mass. But that bound has been slashed nearly in half, from 0.8 eV in the previous analysis,[1] released three years ago, to 0.45 eV in the current one.[2]

Other methods of probing the neutrino mass have, at least ostensibly, generated more stringent bounds. But they rely on theoretical assumptions that might not be correct. KATRIN's kinematic approach, in contrast, assumes nothing beyond the conservation of energy and momentum.

Flavor change

The kinematics of beta decay—in which a neutron emits an electron and transforms into a proton—are what led researchers to learn that neutrinos exist in the first place. If the electron and proton were the only products of the decay, experimental observations would be inconsistent with the law of conservation of energy: The energy imparted to the proton and electron is not always the same, whereas the energy released in the decay should be.

The discrepancy led Wolfgang Pauli to propose in 1930 that beta decay also produced a third, undetected particle that carried the rest of the kinetic energy. The prediction was confirmed in the 1950s by the first direct detection of neutrinos, which was finally honored with a Nobel Prize in 1995 (see Physics Today, December 1995, page 17).

According to kinematic analyses of the energy and momentum released in beta decay, the neutrino mass was (and still is) indistinguishable from zero: Its momentum is directly proportional to its kinetic energy, as expected for a massless

FIGURE 1. TO SEEK OUT SIGNS of neutrino masses, the KATRIN (Karlsruhe Tritium Neutrino) experiment collects and analyzes the electrons emitted in the beta decay of tritium. Successfully wrangling the electrons requires a spectrometer, the inside of which is shown here, that's the size of a house, the shape of a blimp, and permeated with a magnetic field. (Photo courtesy of M. Zacher/KATRIN Collaboration.)

particle traveling at the speed of light. But the mass must be nonzero because neutrinos undergo the unusual phenomenon of flavor oscillation.

Neutrinos come in three flavors that correspond to the three types of charged lepton: electron, muon, and tau. The particle reactions that create and consume neutrinos all conserve lepton family number. For example, beta decays, which produce electrons, also produce electron antineutrinos. And in the Sun and other stars, nuclear fusion reactions, which convert some protons into neutrons and consume electrons in the process, produce vast numbers of electron neutrinos.

But by the time solar neutrinos reach Earth, they're not all electron neutrinos anymore: Some have transformed into muon neutrinos or tau neutrinos. Figuring out that that was what was going on—and that there was nothing wrong with models of the Sun or with neutrino-detection experiments—was a multidecade quest that culminated in a 2015 Nobel Prize (see Physics Today, December 2015, page 16).

Flavor oscillation is not like other particle transformations. It's not the result of any of the four fundamental forces; rather, it's due to the nature of neutrino states. Each flavor state—electron, muon, and tau—is a superposition of three mass states, m_1, m_2, and m_3. As a neutrino propagates through space, the relative phases in the superposition change, which introduces a probability that the neutrino can be detected as a different flavor. Because flavor oscillation requires that the masses all be different, at least two of them must be nonzero.

Measuring mass

Flavor-oscillation measurements reveal relationships between the three masses—specifically, the differences between their squares—but not their absolute values. If $m_1 = 0$, then m_2 is about 0.01 eV and m_3 is about 0.05 eV. But m_1 could be considerably larger than that, in which case the other masses are significantly larger too.

That's where KATRIN, which started collecting data in 2019, comes in. In a return to the roots of neutrino physics, the KATRIN researchers strive to learn something about neutrino masses from the distribution of kinetic energies produced when tritium decays into helium-3. The neutrino, which always goes undetected, carries some amount of kinetic energy. Of the rest, the vast majority is carried by the electron, with a small fraction taken by the much heavier ^{3}He nucleus.

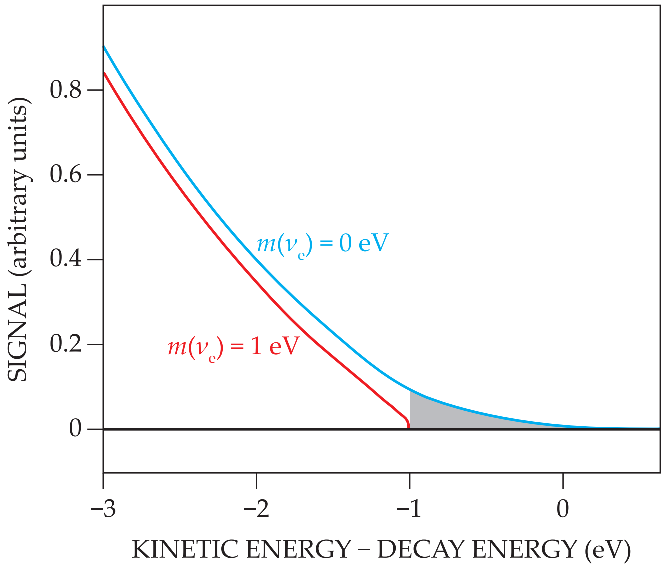


FIGURE 2. ZERO AND NONZERO VALUES of the electron neutrino mass $m(\nu_e)$ can be distinguished, at least in principle, based on the shape of the spectrum of kinetic energies released in tritium beta decay. (Image courtesy of the KATRIN Collaboration.)

The principle of the experiment is illustrated in figure 2. (See also Physics Today, August 2017, page 26.) If the electron neutrino had no mass, the measured energy distribution would look like the blue curve in the plot, with the maximum kinetic energy exactly equal to the energy released in the beta decay. But because the neutrino does have mass, the distribution must look something like the red curve, with the maximum kinetic energy falling short of the total decay energy.

The analysis, however, is not quite that simple. The blue curve is hypothetical, not something that can actually be measured. So the researchers can't derive the neutrino mass from the width of the gray shaded region because there's no gray shaded region to observe. Instead, they aim to infer the mass from the curve's shape: elongated and concave up for zero mass, or slightly snub and concave down for nonzero mass.

Collecting enough data to make that inference is no easy feat. The researchers are reliant on the rare beta decays—fewer than one in a trillion—in which almost all the decay energy is carried by the electron and almost none by the neutrino, and they must catch those rare energetic electrons and measure their energies precisely. And they need to avoid sources of background, such as stray electrons and spurious detector signals, that skew the shape of the observed curve and garble the mass measurement.

Some garbling, however, is inevitable. When the researchers fitted their data for the square of the electron neutrino mass, they got a negative number—a physical and mathematical impossibility but reflective of the fact that, due to background events, the observed energy curve was actually even more elongated than would be expected in the zero-mass case.

The question then becomes: How large can the neutrino mass be and still be consistent with the KATRIN data? Pinning down an answer was a painstaking and time-consuming quest to understand all the sources of background and how they affect the observations. The analysis took years—the new paper, submitted in 2024 and published in 2025, includes data only through 2021. "We need to be very cautious," says Alexey Lokhov of the Karlsruhe Institute of Technology, "because we're entering a region of parameter space that no one has measured directly and no one will be able to cross-check for a couple of years at least." In the end, the researchers concluded at the 90% confidence level that the electron neutrino mass must be less than 0.45 eV.

Not like the other particles

What does it mean that the electron neutrino mass—really, the weighted average of m_1, m_2, and m_3 that creates the electron neutrino flavor state—is less than

0.45 eV? The specific implications are not yet clear, but the first striking feature is how anomalously small the mass is.

The typical mass scale in particle physics is the gigaelectron volt, or 1 billion eV. Protons and neutrons each weigh in at about 1 GeV, and the Higgs boson at 125 GeV. Even the electron, the next lightest particle after the neutrinos, has a mass of 0.5 MeV—more than a million times that of the neutrino. In other words, the KATRIN researchers point out, the electron neutrino is to the electron as a 1 kg bag of sugar is to five blue whales.

The mass imbalance between neutrinos and other particles has been known for years, even before KATRIN provided any constraints on the specifics. It could have to do with the mechanism or mechanisms that endow particles with their mass. Most particles get their mass through coupling to the Higgs field (see Physics Today, September 2012, page 12 and page 14). But the standard model says that the Higgs field shouldn't couple to neutrinos—which is why it says that neutrino masses should be zero.

Could neutrinos acquire their mass from some other mechanism, and if so, what? The answers are necessarily speculative, but one plausible theory proposes what's called a seesaw mechanism, which creates both a set of neutrinos with anomalously small masses and a set of counterparts with anomalously large ones. (See the Quick Study by Rabi Mohapatra, Physics Today, April 2010, page 68.) Although the high-mass particles don't show up in the universe today, they could have affected the dynamics of the early universe. If so, then knowing just how anomalously small the neutrino masses are is a step toward understanding their massive counterparts.

Neutrinos' effect on cosmology doesn't stop there. Because ordinary low-mass neutrinos are so numerous, even in the most remote parts of intergalactic space, they can have a significant effect on the expansion of the universe. Relativistic neutrinos in an expanding universe behave like radiation: Their number density is diluted and their wavelengths are stretched over time. But as the expansion continues, they behave more like matter, which has a cosmologically different effect. When the transition occurs depends on their mass.

And cosmology also offers an alternative route to measuring the neutrino mass. Neutrinos pervade the universe and imbue all space with a background mass density: Not only do they not clump together to form large-scale structures, but their gravitational influence on other matter would have slowed its own clumping. In a sufficiently detailed comparison of structure in the early universe and today, the effect should be measurable.

But last year's results from the Dark Energy Spectroscopic Instrument (DESI) at the Kitt Peak National Observatory show no sign of such a neutrino-mass effect.[3] The DESI researchers estimated an upper bound on $m_1 + m_2 + m_3$ of 0.072 eV. That's at least ostensibly more than an order of magnitude tighter than KATRIN's result and quite close to the lower bound from flavor oscillation (from which $m_1 + m_2 + m_3 > 0.06$ eV). "It's starting to seem worrisome that cosmology doesn't show any hint of the neutrino mass," says Christoph Wiesinger, a KATRIN team member from the Technical University of Munich. "It should have seen something by now."

But DESI's analysis assumes the standard cosmological model, in which dark energy takes the form of a cosmological constant that doesn't change over time or space. And DESI's results this year cast doubt on that assumption.[4] If the density of dark energy is allowed to evolve over time, the sum of the neutrino masses could be much larger and still be consistent with the DESI observations.

"A terrestrial mass measurement would be helpful for cosmology," says Lokhov. "It might help to fix a parameter in their model." KATRIN's kinematic data-taking is scheduled to wrap up later this year, for a total of four years' worth of data not included in the latest analysis. That should be enough to shrink the bound on the neutrino mass by another factor of two.

Johanna L. Miller

References

1. KATRIN Collaboration, *Nat. Phys.* **18**, 160 (2022).
2. KATRIN Collaboration, *Science* **388**, 180 (2025).
3. DESI Collaboration, https://arxiv.org/abs/2404.03002.
4. W. Elbers et al. (DESI Collaboration), https://arxiv.org/abs/2503.14744.

A drone helps make a challenging Greenland glacier measurement

A rare investigation of a marine-terminating glacier in the Greenland winter yields evidence of melting at the base of the glacier.

Eqalorutsit Kangilliit Sermiat is a vast glacier that lies on bedrock below sea level and stretches into the Arctic waters of Sermilik Fjord off the coast of southern Greenland. When the fronts of such marine-terminating glaciers melt at depth, the meltwater rises through the salty seawater and triggers mixing that affects, among other things, the melting rates of the glacier front and the availability of nutrients for marine life.

The melting of Eqalorutsit Kangilliit Sermiat and other Greenland glaciers has been measured in the summer but not during the winter, when fjord ice and icebergs permeate the fjords. The lack of data has left researchers unsure of how to factor wintertime melting into climate models and other studies of Greenland's glaciers and fjords. To help obtain valuable winter data, Nanna Karlsson of the Geological Survey of Denmark and Greenland in Copenhagen and her colleagues fitted a commercial drone helicopter, shown in figure 1, with a deployable sensor for measuring water salinity, temperature, and depth.

In March 2023, Karlsson and her Geological Survey colleagues, along with researchers from Aarhus University in Denmark and the Greenland Institute of Natural Resources, navigated the roughly $15 000 custom vehicle out into icy waters several kilometers from the front of Eqalorutsit Kangilliit Sermiat and lowered the sensor as deep as 100 m underwater. The researchers obtained measurements closer to the glacier front by manually drilling through the frozen fjord surface.

By deploying sensors both beside the glacier front and farther out in the fjord, the researchers were able to compare water properties and determine the potential influence of glacial meltwater. As shown in figure 2, the drilling yielded evidence that Eqalorutsit Kangilliit Sermiat was melting at depth, though the amount of meltwater was considerably less than the amounts measured in the summer. The water measured by the drone was warmer and saltier with increasing distance from the glacier, indicating a waning influence of glacial discharge. The findings imply that the meltwater rises and gets trapped beneath the icy fjord surface near the glacier front.

FIGURE 1. AN UNCREWED HELICOPTER hovers over Sermilik Fjord in March 2023 as a sensor that was lowered under the surface collects data. (Image adapted from K. Hansen et al., *Nat. Geosci.* **18**, 219, 2025.)

Karlsson and colleagues propose that the meltwater originates at the glacier's base, where ice grinds against bedrock. The winter discharge could affect the conditions of the nearby icy waters and facilitate faster melting rates in the future. Wintertime melting may also nudge nutrients up toward the surface and fuel the phytoplankton bloom that occurs in the area each spring. Future measurements, perhaps aided by drones, could help quantify the climatic, ecological, and other effects of the wintertime melting of glaciers like Eqalorutsit Kangilliit Sermiat. (K. Hansen et al., *Nat. Geosci.* **18**, 219, 2025.)

Andrew Grant

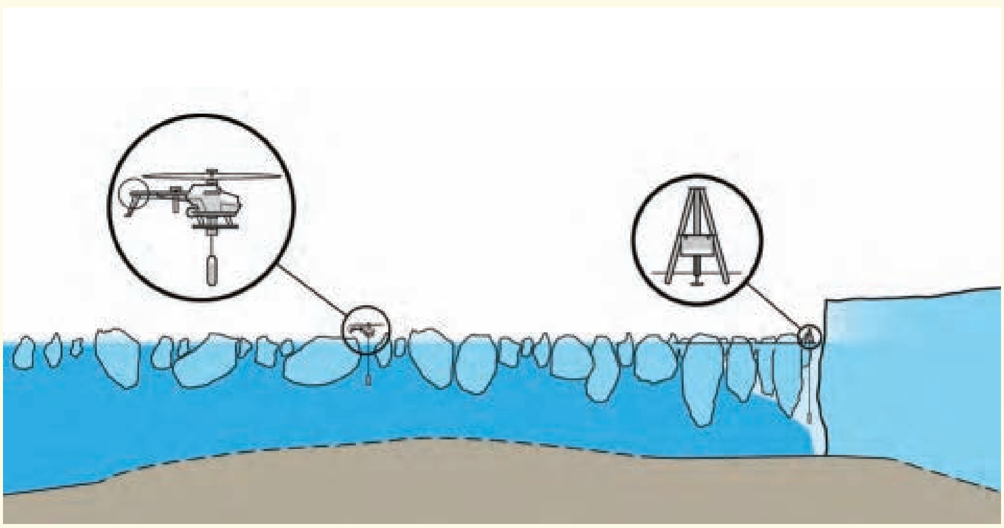


FIGURE 2. WATER MEASUREMENTS made via drill close to the glacier front and via drone farther out to sea indicate the existence of a supply of relatively fresh water (light blue beneath the drill) near the glacier. (Image adapted from K. Hansen et al., *Nat. Geosci.* **18**, 219, 2025.)

A hovering resonator enhances the quantum Hall effect

Electromagnetic vacuum-field fluctuations are capable of reshaping the electronic correlations of solid-state systems.

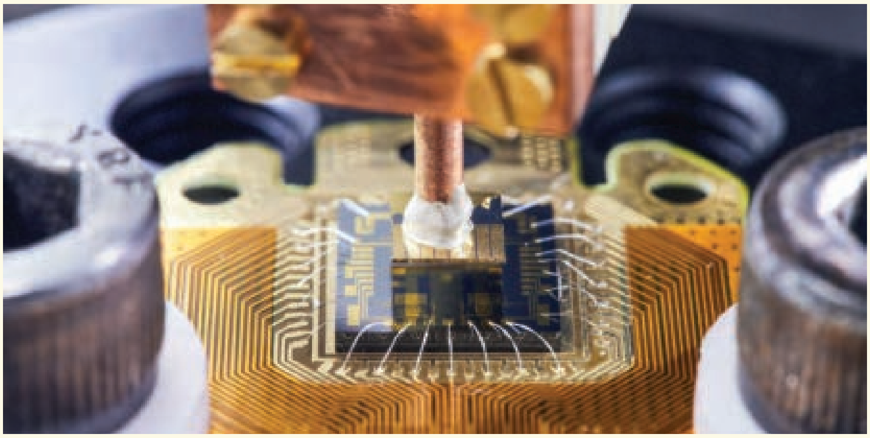


FIGURE 1. THE GOLD SQUARE attached to the end of the copper post contains a split-ring resonator that influences the ground-state electromagnetic fields in its vicinity. By raising and lowering the resonator above a sample chip, researchers control how the sample exhibits the quantum Hall effect. (Photo by Kilian Kessler/ETH Zürich.)

The eerie music of a theremin—the instrument that's played without touching it—seemingly comes from thin air. Eerier still is a new result from Jérôme Faist (of ETH Zürich in Switzerland) and colleagues. By holding a metallic resonator (the gold square attached to the copper post in figure 1) above an electronic chip, the researchers enhance how the chip exhibits the quantum Hall effect. But whereas a theremin's arms are active antennas that send and receive RF signals, neither the resonator nor the chip does the same. Instead, the effect stems entirely from the ground state of the electromagnetic vacuum.

Even the darkest, emptiest vacuum isn't devoid of electromagnetic fields. For the same reason that a particle can't perfectly settle at the bottom of a bowl-shaped potential—because to do so would violate Heisenberg's uncertainty principle—the energy in any of a system's electromagnetic modes can never be exactly zero. But the set of allowed modes can be restricted through the placement of conductive materials, which impose electromagnetic boundary conditions. One manifestation of that is the Casimir effect, a measurable attraction between two conductors that stems from the diminished electromagnetic vacuum energy in the space between them (see the article by Alex Stange, David Campbell, and David Bishop, Physics Today, January 2021, page 42).

Can vacuum fields also be used to manipulate the exotic many-body states of a condensed-matter system? It should be possible, and a 20-year-old theory paper by Cristiano Ciuti (also a coauthor on Faist and company's new paper) and colleagues drew a road map for identifying systems in which the vacuum-field energy scale could be large enough to have a measurable effect. But experimental searches have struggled until now, in part because electronic states are often probed optically, so the measurements are less sensitive to the vacuum field's effect on the ground state. The quantum Hall effect is a fortunate alternative: It's probed entirely through measurements of charge transport.

In the classical Hall effect, a magnetic field deflects an electric current to create a voltage perpendicular to them both; the Hall resistance—the transverse voltage divided by the current—scales linearly with the magnetic field. In the quantum Hall effect, plateaus appear at conductance values that correspond to integer multiples of e^2/h, where e is the electron charge and h is Planck's constant. In the still more exotic fractional quantum Hall effect, conductance plateaus appear at fractional multiples as well. Figure 2 shows some of Faist and colleagues' results: Bringing the resonator within tens of microns of the chip (shown in the plot by progressively lighter blue curves) enhances the flatness of the 5/3 quantum Hall state.

The theory behind the enhancement mechanism, like a lot of solid-state theory, is complicated and not well understood. But there are hints of promising directions for future development. The fractional quantum Hall effect can be understood in terms of the magnetic field coupling to the electrons and forming them into quasiparticles with nonintegral charge. If the resonator's influence is to assist with the quasiparticle formation, it's at least plausible that manipulating the vacuum fields in other solid-state systems could help drive the assembly of other useful quasiparticles, such as the Cooper pairs that are responsible for superconductivity. (J. Enkner et al., *Nature* **641**, 884, 2025.)

Johanna L. Miller

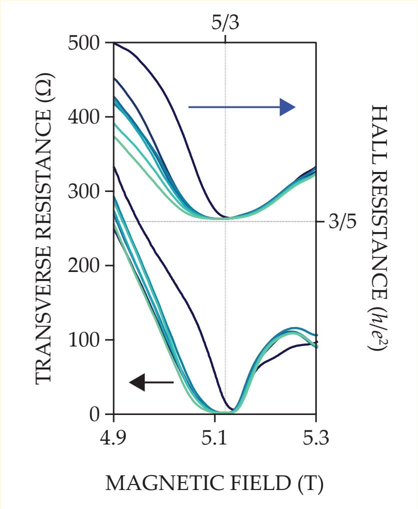


FIGURE 2. THE FRACTIONAL QUANTUM HALL EFFECT is characterized by flat-bottomed curves for both the transverse and Hall resistances. When the resonator is too far from the sample to have an effect (darkest blue curves), the observed resistances barely qualify. But as the resonator is brought closer to the sample (lighter blue curves), the curves become flatter: The fractional quantum Hall effect is enhanced. (Figure adapted from J. Enkner et al., *Nature* **641**, 884, 2025.)

The story of Mars's early atmosphere is told in carbonate rocks

Newly found iron-rich minerals suggest that carbon dioxide cycled in and out of the planet's ancient atmosphere.

Mars's atmosphere, compared with Earth's, is thin and wispy. Some of it was lost to space after the planet's magnetosphere failed some 4 billion years ago (see the Quick Study by Bruce Jakosky, Physics Today, April 2022, page 62). The current air pressure—only 0.6% that of Earth's—ensures that the Martian surface stays cold and dry. Mars's current atmosphere has roughly 6 millibars of carbon dioxide, but if the past atmosphere had at least tens of millibars of CO_2, as models suggest, the climate could have been warm enough to support liquid water (see the article by Bruce Jakosky and Michael Mellon, Physics Today, April 2004, page 71).

For several decades, researchers studying Mars have used rovers and orbiters to search its sediments for evidence of CO_2. The gas, after being emitted from volcanic eruptions, may have reacted with liquid water to form sedimentary carbonate rocks. Yet despite many remote-sensing observations, not enough carbonates have been found on the Martian surface to support an atmosphere rich with CO_2. Now Benjamin Tutolo (University of Calgary) and colleagues have found a large carbonate reserve in the subsurface, in an iron-rich mineral not seen before on Mars. Called siderite, the mineral was found to have a nearly pure iron composition.

In 2022, NASA's *Curiosity* rover reached Mount Sharp, a 5-kilometer-tall stack of sediments in Gale Crater, which is thought to have once been a lake. There, the rover drilled four samples. Tutolo and colleagues' analysis identified not just siderite but also several other minerals, including sulfates and iron oxides. The finding of all the minerals together suggests that siderite, with its especially pure composition, precipitated out of an aqueous solution underground. Later, infiltrating fluids partially dissolved the siderite to form iron oxides. If the released carbon made its way to the atmosphere, it would have contributed to a Martian carbon cycle and potentially prolonged the duration of liquid water on the surface because of continued greenhouse warming.

Tutolo and colleagues speculate that the sulfate minerals masked the siderite from the reflectance and thermal emission spectroscopy approaches of remote-sensing instruments. Various sources would have added CO_2 to Mars's ancient atmosphere. Using a broad estimate for the amount of siderite that may be in the planet's subsurface, the researchers predict that the carbonate-rich sediments contributed 2.6–36 millibars of CO_2 to the total. The higher end of the range would offer enough atmospheric pressure to stabilize liquid water on the surface. But more CO_2—potentially lost through other means such as atmospheric escape—would have been required to raise temperatures to habitable conditions.

The finding of siderite and its partial dissolution suggests a planetary carbon cycle. When atmospheric CO_2 was plentiful, some was sequestered to form siderite, and then subsequent rock–water interactions released some CO_2 back into the atmosphere. Now that the mineral has been found with sulfates, reanalysis of existing remote-sensing data could help determine whether large siderite deposits exist elsewhere. On the ground, other Mars rovers—*Perseverance* and *Spirit*, before it lost contact with Earth—spotted iron-rich oxides close to sulfates in various locations. The samples that *Perseverance* has collected, and that NASA hopes to bring back by 2040, will likely yield new clues about the planet's ancient atmosphere. (B. M. Tutolo et al., *Science* **388**, 292, 2025.)

Alex Lopatka PT

PILES OF ROCKS, which could have formed from floods of water, are seen in this 180° panorama of a channel in Gale Crater, captured by NASA's *Curiosity* Mars rover on 31 March 2024. (Image by NASA/JPL-CaltechMSSS.)

Funding uncertainties muddle graduate admissions

Aiming to bring on PhD students who they can keep commitments to, universities are adjusting their admissions processes and offers.

SARA EARNEST (LEFT) AND HUDSON LAZZARA (RIGHT) are among the physics PhD applicants who got caught up in a chaotic, confusing admissions process this year. (Photos by Otis Michael Jr and Chase Hayes, respectively.)

"Strange and harrowing." That's how Sara Earnest describes the process of applying for physics PhD programs this year. She graduated in May from Johns Hopkins University with two and a half years of undergraduate research experience. But just two weeks shy of the 15 April national deadline for prospective students to commit to graduate programs, she had been wait-listed by one, rejected by seven, and was still waiting to hear from three.

In the end, Earnest didn't get into any of them. She plans to try again next year. In the meantime, she's headed to Chile for a research stint studying black holes.

Hudson Lazzara didn't get into any of the five physics graduate programs he applied to his first go-round; he graduated in 2024 from California Polytechnic State University. This year, he applied to 13 programs at 9 schools and got into 2: complex systems at Arizona State University and physics at the University of Oregon.

But "getting in" turned out to be a slippery concept. He visited Arizona State in early April and was told he would get a one-year fellowship, but by the end of the month, the paperwork hadn't come through. Oregon accepted him in January, but instead of following up with a formal offer, the university wait-listed him. On 17 April, he got the news that he was in with five years of guaranteed support. He starts there this fall.

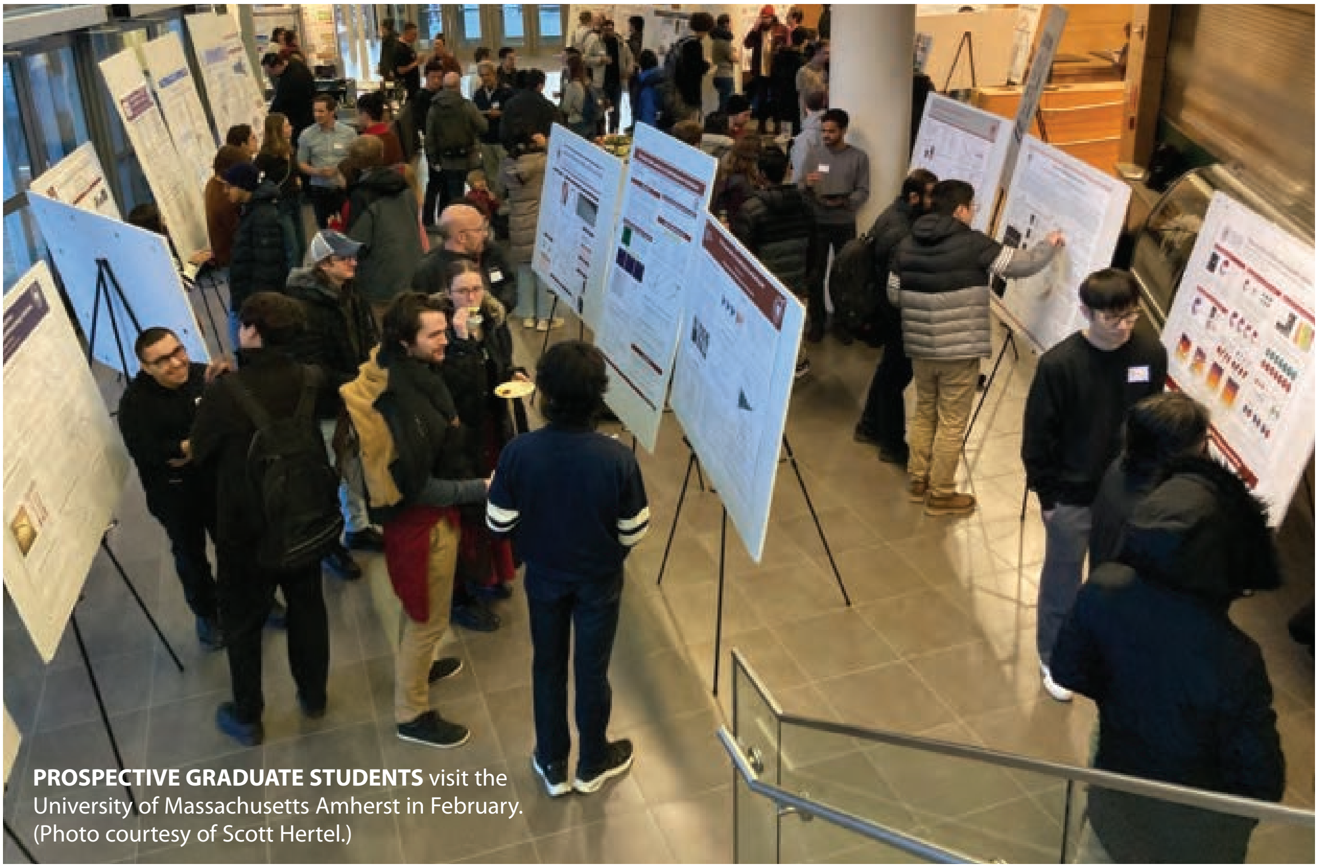
Earnest's and Lazzara's experiences reflect both a long-term trend of growing competition to get into US physics PhD programs and the disruptions to higher education since the start of the current Trump administration, in particular the prospect of severe reductions in funding. (See "Rapid-fire changes in federal funding stoke uncertainty in US universities," Physics Today online, 14 February 2025.)

In April, the chairs of US physics and astronomy programs expected to accept 13% fewer PhD students—about 600—for fall 2025 than in the previous year, according to a survey conducted by the American Institute of Physics' statistics team (see Physics Today, June 2025, page 23; AIP publishes Physics Today). The decline in first-year enrollment was anticipated to be 25% at private institutions and 7% at public ones.

For many graduate schools, the admissions process this year was fraught: Would applicants be more likely to accept offers knowing that other schools were reducing the sizes of their incoming classes? Would departments be able to keep their commitments to incoming students? Would they be able to continue supporting their senior graduate students?

Reactions to unknowns

In the fall and winter, physics graduate admissions committees typically poll their faculty members to see how many want to take new PhD students and how many expect to be able to support them on external grants. The committees also look at their institution's need for teaching assistants (TAs)—it's common for physics PhD students to work as TAs for their first year or two and thereafter as research assistants (RAs) on their adviser's grants. Many admissions commit-



PROSPECTIVE GRADUATE STUDENTS visit the University of Massachusetts Amherst in February. (Photo courtesy of Scott Hertel.)

tees try to align their offers such that prospective students' research interests roughly match those of the groups looking to expand. Then they make their best guess as to how many offers to make to end up with their target class size.

This year, however, historical norms for making those guesses were out the window. Some departments reduced the number of offers that they made but still got large numbers of acceptances; others introduced wait-lists, tiered offers, or other measures to try to control their final numbers.

Last year's incoming class of 75 physics PhD students at the University of Illinois Urbana-Champaign was unusually large, so the department was aiming to scale back to its usual 45–50 students even before the presidential elections, says Lance Cooper, associate head for graduate programs in physics. "We reduced the number of offers by about one-third," he says. "We nervously watched as the numbers rolled in. In spite of our best efforts, we got 78 accepts."

"We went back and forth about what our strategy would be," says Steven Rolston, chair of physics at the University of Maryland (UMD) in College Park. "We made fewer offers, and, for the first time, they were tiered." The strongest applicants received the department's standard offer of two years of guaranteed support as a TA, he says. A second set of applicants was offered one year of guaranteed TA support, and a third tier was accepted on the condition that they find a faculty member who would hire them as an RA from the get-go. UMD physics gets 1000 applications for 45 or so spots, says Rolston. This year, they ended up with 36 students.

The physics department at the University of Iowa took a different tack. "We accepted people but waited to send out employment offers," says physics and astronomy chair Mary Hall Reno. "We were working on the basis of what's in hand, not on the basis of what is likely to be funded because we just don't know anymore." In the end, she says, 10 students will start in the fall, 8 as TAs, 1 as an RA, and 1 with a fellowship. With fewer graduate students than usual available to serve as TAs, she adds, "we will have to squeeze more undergraduates into lab sections."

By the time the physics department at the University of Oregon began making offers, the university administration was gun-shy: The chemistry department had so many students wanting to come that it ended up rescinding offers to applicants who had not yet accepted, says physics chair Richard Taylor. The experience prompted the university administration to insist that the physics department "drill down" to figure out how many students it could support. That included estimates about when senior graduate students would complete their degrees and internal pressures related to tuition dollars—resulting in the university allotting fewer teaching assistantships. It also took into account pessimism about federal funding.

"We came down to six," Taylor says. The department's incoming class usually numbers around 18. "The good news is the university allowed us to offer guaranteed support for five years." To limit to six incoming PhD students, the department introduced a waiting list—the one that Lazzara had been on—and made offers as space allowed. Taylor says he expects to use wait-listing going forward.

Compounded implications

Other departments similarly staggered their offers, made fewer of them and offered less, conditioned them on student performance and availability of funds, and held their breath. For many departments, the warning signs came too late in the admissions cycle to respond. "We didn't do anything out of the ordinary," says Dan Hooper, a member of the graduate admissions committee at the University of Wisconsin–Madison. "But if next year it's clear we don't have money to sustain students, we would have no choice but to admit fewer students."

Faculty members also worry about senior graduate students, who often do not have guaranteed funding but traditionally have been covered by research grants. Historically, if an advanced graduate student lost their research assistantship, they could often fall back on a teaching assistantship. But institutions would not be able to step in like that for large numbers of students. "We have told our faculty that they should encourage their students to graduate faster," says UMD's Rolston.

On top of worries about funding are concerns about the US government's policies toward foreigners. In physics, international students have long made up a huge chunk of the graduate body. So far, at least, they still want to come to the US: At the University of Illinois Urbana-Champaign, for example, where the domestic-to-international ratio is usually about 50:50, for the 2025 incoming class, it's 44:56. But recent developments, including US government threats to revoke visas of Chinese students, a pause on conducting interviews for student visas, and the spate of detainments of people from multiple countries, could make it harder for some to get into the US and make the country a less desirable destination. On the flip side, some US professors are advising their graduating seniors to apply for PhD spots abroad.

In the long term, shrinking graduate enrollment would damage both individual departments and the broader physics community. "Every step along this trajectory will be harder to advance, and there will be far fewer opportunities for talented young researchers to make a career," says Hooper. "The US science system is the envy of the world, and we should think twice before we decide to tear it down."

Toni Feder

Scientists scramble to save threatened federal research databases

Amid funding and workforce cuts, US physical sciences databases are in jeopardy.

Assistant professor Patrick Rafter was sitting in his University of South Florida office in April when he saw an unusual banner across his computer screen. The federal database he had accessed for his research that day would be removed in a matter of weeks. Rafter had visited the NOAA-maintained Index to Marine and Lacustrine Geological Samples (IMLGS) regularly over his 20-year career studying Earth's climate.

An emergency video call soon followed. Rafter listened as NOAA staff spoke with university research leaders about finding a new home for the database, which served as a browsable repository for all lake and marine sediment cores kept at institutions around the world. Listed in the database were scientific samples worth "millions and millions of dollars," says Val Stanley, the Antarctic core curator at Oregon State University's repository. Without it, scientists would need to individually contact the nearly 30 global repositories when looking for cores.

The IMLGS is one of at least 25 databases and products that NOAA has announced for retirement since April. Some of the databases would still be available online, but no longer updated, while others would be removed from public websites and available only upon archive request. The databases include historic earthquake recordings, satellite readings of cloud radiative properties, and a tool for studying billion-dollar disasters. Rick Spinrad, former NOAA administrator under Joe Biden, says that climate-related physical sciences data seem particularly at risk.

Still more data could be taken offline or no longer updated because of proposed cuts in President Trump's fiscal year 2026 budget request, which would significantly reduce or cut science programs at NSF, NASA, the Environmental Protection Agency, the National Park Service, NOAA, and the Department of Agriculture. Experts around the world use the data in myriad ways—for example, scientists use them for their research; companies, for developing products; lawmakers, for crafting legislation; and nongovernmental organizations and nonprofits, for creating and improving community services.

Grassroots data rescue efforts to protect such vital resources began in Trump's

SEDIMENT CORES at the US Geological Survey's Pacific Coastal and Marine Science Center in Santa Cruz, California, and other repositories around the world are no longer searchable online after NOAA removed its public core database in May. (Photo by Rex Sanders, USGS Pacific Coastal and Marine Science Center.)

first term (see Physics Today, March 2017, page 31). Those initiatives have continued and expanded. The volunteer-led Data Rescue Project, the nonprofit Public Environmental Data Project, and groups housed at universities have downloaded more than a thousand federal public datasets, webpages, and online tools.

But backing up data for safekeeping is often only the first step. Ideally, datasets would have dedicated staff to curate them and to support continued data collection.

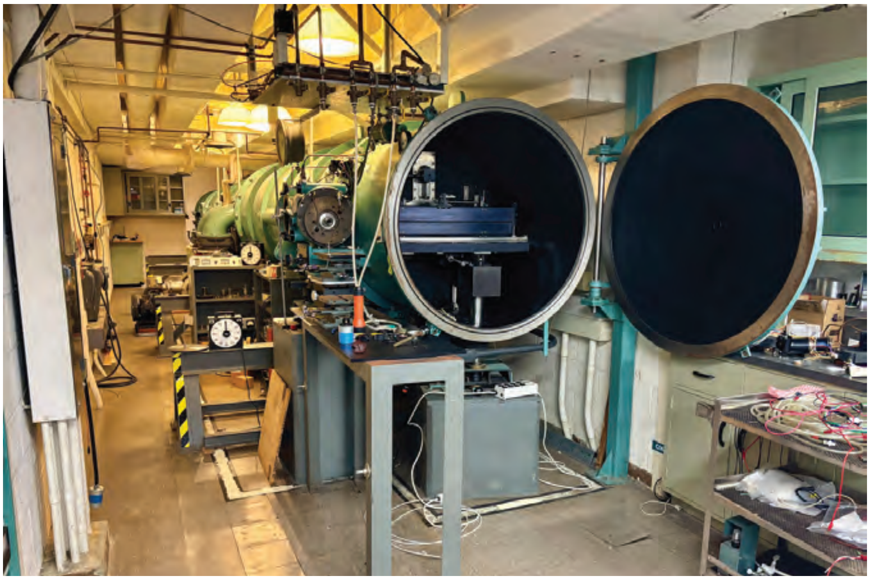
Finding new guardians

Yuri Ralchenko sounded the alarm bells in mid-March. The atomic spectroscopy group he led at NIST was being shut down ahead of potential reduction-in-force staff cuts in the Department of Commerce, which oversees the institute. Said Ralchenko in an email to the scientific community, "The very first scientific paper from the National Bureau of Standards [NIST's original name] back in 1904 was on spectra of mixed gases. . . . Unfortunately, the story of atomic spectroscopy at NIST is coming to an end."

The group oversaw the Atomic Spectra Database, a collection of critically evaluated reference spectra of neutral and ionized atoms. Specialists in medical fields, semiconductor chips, astronomy, and other areas used the database; it received about 70 000 search requests per month. If the team of six technical staff lost their jobs, the database would shutter. Hundreds of emails poured in, and a change.org petition racked up almost 5600 signatures. At a time when tens of thousands of federal workers were losing their jobs, the six-person group attracted a large outpouring of support.

The outreach paid off. The University of Maryland, through a collaboration with NASA Goddard Space Flight Center, agreed to employ the team. But the transition will take time, Ralchenko says, because the researchers must move laboratory equipment and install it in their new laboratory housed at Goddard. The Atomic Spectra Database will still live on NIST's website, but it won't receive any new updates. By the end of the year, the team plans to release a mirrored version of the data repository at Goddard, likely under a new name, that will prioritize astronomy and astrophysics applications going forward.

The team's instrumental research program that provides some of the information for the database will reach full operations in 1.5 years, says Ralchenko. "Since we will be busy with the move, we'll not be able to produce new results in the meantime."



A SPECTROGRAPH from NIST's former atomic spectroscopy group will not be relocated to the team's new home at NASA's Goddard Space Flight Center, since it is too large and delicate to move. (Photo courtesy of Yuri Ralchenko.)

For the IMLGS, salvation also came rather quickly. About two weeks after the initial group video call, a new home was announced: The NSF-funded SESAR2 (System for Earth and Extraterrestrial Sample Registration), an online service hosted at the Lamont–Doherty Earth Observatory of Columbia University. SESAR2 went to work creating a mirror of the frozen NOAA database, but their plans for allowing future additions were listed as "TBD."

University of South Florida's Rafter was glad to see that IMLGS would live on, but he couldn't wrap his head around why the change was necessary. "Public data should be in the public domain," he says. "And run by the feds."

Future data at risk

Brittany Janis, executive director of the nonprofit Open Environmental Data Project, says that some physical sciences data may stay accessible but frozen in time.

C. David Keeling began sampling atmospheric carbon dioxide levels in 1958 at Mauna Loa Observatory in Hawaii. The data record, now called the Keeling curve, shows average CO_2 concentrations increasing from roughly 313 parts per million in the first measurement to nearly 430 ppm by April 2025. It has served as one of the key indicators of climate change.

Keeling died in 2005, and his son, Ralph Keeling, took up the work as a professor at Scripps Institution of Oceanography at the University of California, San Diego. Ralph is not worried that the data could be lost because they have a relatively small digital footprint. But he does fear for the future of the data record: Trump's FY 2026 budget request would cut NOAA's programs that support taking measurements for the Keeling curve.

The agency also maintains air measurements at more than 50 stations around the world and provides calibrated CO_2 samples to hundreds of groups conducting their own measurements, says Keeling. When considering the full scope of federally funded climate- and environment-related observations, he says, "there's no way that private philanthropy or other organizations can take over more than a fraction of what's going on."

Janis says that if the federal government abdicates its role in collecting and hosting data, private companies may selectively pick up the slack—but the products could be proprietary and thus out of reach from the research community and the public.

Jenessa Duncombe

IN 1983, SALLY RIDE became the first US woman to go to space. (Photo from NASA.)

Q&A: Tam O'Shaughnessy honors Sally Ride's courage and character

In a new documentary, Ride's life partner of 27 years chronicles the astronaut's leadership, resilience, and dedication to science.

Physicist Sally Ride made history in 1983 as the first American woman to go to space. In the decades after, Ride devoted herself to teaching physics and advocating for women in STEM. Out of the public eye, Ride also built a happy life with her partner, Tam O'Shaughnessy. Ride's journey is told by O'Shaughnessy and others close to Ride in the National Geographic documentary film *SALLY*, which started streaming 17 June on Disney+ and Hulu.

A natural athlete, Ride grew up in Southern California and became a nationally ranked tennis player. She considered playing professionally but chose to attend Stanford University to study physics and English. She continued there for her master's and doctorate in physics, intent on being a physics professor, until she saw an article in the student newspaper saying that NASA was hiring astronauts. For the first time, women could apply.

Ride was one of six women chosen for NASA's 1978 class. She went twice to space, where she helped deploy satellites and conduct scientific observations of Earth. After her time at NASA, Ride worked as a physics professor for 18 years at the University of California, San Diego (UCSD), where she taught, mentored students, and studied free-electron lasers. During that time, O'Shaughnessy, Ride, and their friends launched the science education company Sally Ride Science. Ride passed away from pancreatic cancer in 2012 at age 61.

O'Shaughnessy is a writer and executive producer for *SALLY*, and she is the executive director of Sally Ride Science at UCSD. Physics Today spoke with her about Ride's life and legacy.

PT: What do you want people to take away from watching *SALLY*?

O'SHAUGHNESSY: Most people think of Sally as the astronaut: the first American woman to go to space. She was so much more than that. She was an athlete, an excellent tennis player, a science writer, and an entrepreneur. But the thing that was most important to her was being a physicist. Sally was an astronaut for nine years, and she was a physicist for most of her life.

I think the other fuller picture of who Sally was includes our relationship. Because of the culture around being gay, we knew that we needed to stay quiet about being a couple. Sally Ride Science was dependent on corporate sponsorships, and we thought at that time there was no way that they would back us if they knew that Sally and I were leading the company and were a couple.

PT: How did you two meet?

O'SHAUGHNESSY: We met when we were young at a tennis tournament in Redlands, California, and we stayed friends throughout our junior tennis days. When Sally went to Stanford to study physics, and I went to play on the first-ever women's professional tennis tour, our paths diverged. Then, in 1983, she invited me to come to the space shuttle launch for her first mission.

TAM O'SHAUGHNESSY (LEFT) AND SALLY RIDE, pictured here in Sydney, Australia, in 2004, were life and business partners. Friends since childhood, they authored science books together and ran the company Sally Ride Science. (Photo courtesy of Tam O'Shaughnessy.)

On launch day, I remember very clearly hearing 10, 9, 8, … I could barely breathe. Everyone in the stands was crying, shouting. It was a remarkable experience.

PT: What happened when Sally returned from her first space mission?

O'SHAUGHNESSY: She was so proud to be an astronaut and to be the first American woman to go to space. She was proud of the way she handled her first mission. She just looks so joyous in every image of her in space, so happy and relaxed.

As she started giving talks around the country following her flight, she started feeling anxious and unsettled, and so for the first time in her entire life, she thought, "I need help." She saw a psychologist, who said Sally was an introvert and that Sally should take time after each engagement to recover. She started doing that, and then it got better.

She started accepting speeches in cities where her friends were, and so she visited Atlanta, where I lived, more often. Our friendship felt like we'd never taken a break. We just enjoyed each other so much. In one second, you're best friends, and the next second, it's something romantic. From the first kiss, it felt like this was important.

PT: How did Sally react to the 1986 explosion of the space shuttle *Challenger*?

O'SHAUGHNESSY: Sally was heartbroken. One of her best friends in her astronaut class was Judy Resnik, who was one of the seven crew who lost their lives. Sally was the only woman Ronald Reagan asked to be on the presidential commission investigating the disaster.

PT: How did Sally help uncover the cause of the disaster?

O'SHAUGHNESSY: During the investigation, Sally was passing a NASA engineer in the hall, and he handed her something. He said something like, "You're going to want to read this," and walked away. It was a spreadsheet showing the resilience of the *Challenger*'s O-rings and how pliable they were versus temperature. Sally analyzed the spreadsheet and realized that it was the key to what had happened.

▲ **SALLY RIDE** stands with fellow women astronauts in her 1978 class during a break from training. From left are Ride, Judith Resnik, Anna Fisher, Kathryn Sullivan, and Rhea Seddon. (Photo from NASA.)

▼ **RIDE** communicates with ground control aboard space shuttle *Challenger* during her first trip to space, in 1983. (Photo from NASA.)

Sally thought if she was the one to bring up the spreadsheet during the investigation that people might put together who the NASA engineer was. So she handed the spreadsheet to Major General Donald Kutyna, who was also on the investigation board.

Kutyna hinted about the temperature of the O-rings to fellow commission member and physicist Richard Feynman, who figured it out. Feynman was a showman. At an open commission meeting with the press, Feynman took a glass of very cold water, and put a rubber ring in it, and then pulled it out and it was all brittle. Which was exactly what happened on the *Challenger*.

PT: How did Sally Ride Science come about?

O'SHAUGHNESSY: In 1990, we decided to try our hand at writing children's science books together. Our first book, *Voyager: An Adventure to the Edge of the Solar System*, was very well received. At that same time, Sally and I both noticed articles in the newspaper about how American girls and boys weren't doing well in science or math. We just couldn't understand, because we loved science and math. We learned that the way science is presented across the board persuades kids to think science is boring, done by white men with crazy hair and alone in a lab.

PT: What did you do to challenge those stereotypes?

O'SHAUGHNESSY: We hosted Sally Ride Science festivals around the country at universities. They were one-day events, and kids, their parents, and their teachers could sign up online from a selection of three dozen or so workshops in all sorts of topics, like ocean robotics, photosynthesis, or genetics. The festival had street fairs with local astronomers, veterinarians, and scientists who would have booths with free handouts and activities, such as DNA bracelet making. We also did teacher training on how to engage boys and girls in STEM. And one of the key things was to show role models of contemporary women and men doing science.

PT: In 2013, you accepted the Presidential Medal of Freedom on Sally's behalf at the White House. Can you tell me about that day?

O'SHAUGHNESSY: During the award ceremony, I remember feeling free, honest love and acceptance. My country, the United States of America, was honoring all of whom Sally was and validating our relationship. It sent the message around the world that same-sex relationships are true relationships, worthy of all rights and protections. I only wished Sally could have experienced all of this.

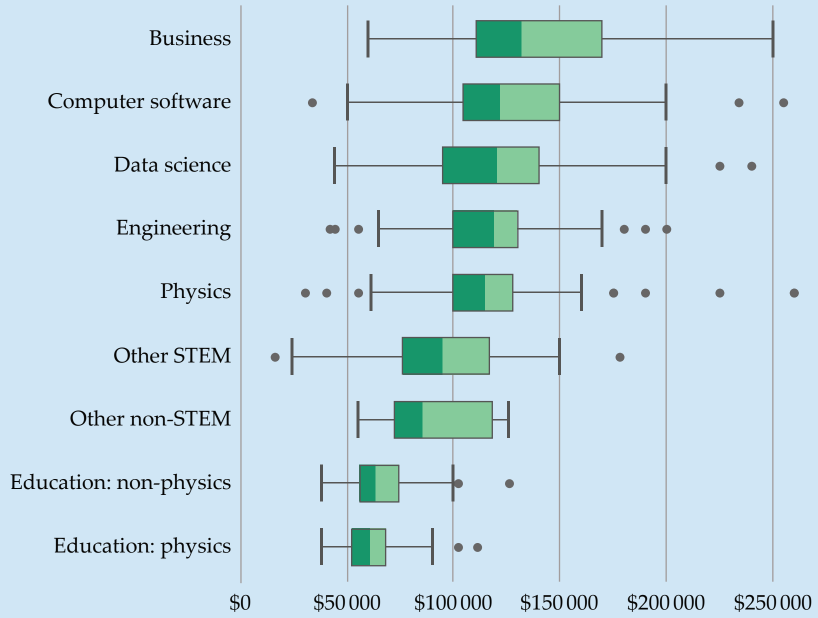
PT: The film pays tribute to Sally's accomplishments but also shows some of her flaws. Can you talk about that editorial decision?

O'SHAUGHNESSY: I was an executive producer, but I didn't make editorial decisions about the film, though I spoke up and added my three cents. The folks at National Geographic didn't want the story to be "everything about Sally is perfect." They wanted the real story.

continued on page 28

Business, STEM jobs offer strong starting salaries to new physics PhDs

Starting salary ranges by field of employment



(Figure adapted from P. Mulvey, J. Pold, J. Tyler, *Who's Hiring Physics PhDs*, AIP Research, 2025.)

The median starting salary for recent physics PhD recipients who secured potentially permanent jobs in computer software, data science, engineering, or physics was at least $115 000 for each of those fields, according to newly published survey data. Nearly three-quarters of new physics PhDs work in those fields.

The data are from a survey of PhD recipients' initial employment for the classes of 2018–22 conducted by the statistical research team of the American Institute of Physics (publisher of Physics Today). As shown in the figure, the highest-earning field was business, with positions offering a median starting salary of $132 000, and the lowest-earning field was physics education, with a median salary of just over $60 000. The US median household income in 2022 was $74 580, per the US Census Bureau.

Survey respondents who were not in physics, engineering, or education were spread across a variety of jobs—in business and accounting, social media, and more. Employers included the health insurance company Aetna, eBay, Facebook, and the Boston Red Sox.

The vast majority of respondents across all fields said they used technical problem-solving skills at least monthly; many said the same for programming, advanced math, and other scientific and technical skills. Among the interpersonal and management skills that survey respondents frequently ranked as required in their jobs were teamwork, technical writing, and managing projects.

Explore employers, salary ranges, and skills needed in different sectors using the interactive charts at https://www.aip.org/statistics/whos-hiring-physics-phds.

Tonya Gary

continued from page 26

PT: Along those lines, fellow astronaut Kathryn Sullivan says in the documentary that Sally may have tried to sabotage one of Sullivan's training exercises by switching off the circuit breakers. What do you think of that?

O'SHAUGHNESSY: I understand Kathryn's uncertainty about what happened. Maybe Sally flipped the switches. But that doesn't strike me as true to who Sally was. I don't know.

PT: What was it like seeing the archival NASA astronaut footage in *SALLY*?

O'SHAUGHNESSY: A lot of that footage is rare and pretty amazing. Seeing Sally's face when she's being dragged through the water in water survival training, that was new. I remember her telling me, "I felt like I was drowning." Members of the press were filming and waiting to ask her questions when she got out of the water.

PT: It's been nearly 50 years since NASA began actively recruiting women and people of color to be astronauts. This year, the Trump administration ended federal diversity, equity, and inclusion [DEI] programs. What do you say to that, and what do you think Sally would say?

O'SHAUGHNESSY: Sally would probably say something similar to what I would say, which is, How moronic. The history of the world and the history of our country are not perfect. Certain people have been second-class citizens—people of color, women, girls. Sally never would have flown to space if NASA hadn't started a DEI-type program. What's wrong with diversity? The world is diverse.

PT: What else can we learn from Sally's life?

O'SHAUGHNESSY: As a person, Sally was one of the most relaxed people I've ever met. She was totally comfortable in her own skin and sure of who she was. I think that's also part of why she didn't need to say, "I'm gay, I'm queer." She lived her life exactly the way she wanted to live it. I think that that is also a major message of the film: Don't let anyone ever tell you what to do with your life or who you should love. And Sally lived that.

Jenessa Duncombe

FYI SCIENCE POLICY BRIEFS

Trump seeks massive cuts to science

Congress is set to consider the steep cuts to science agencies that were proposed in President Trump's budget request for fiscal year 2026, which has drawn outcry from former agency leaders and professional societies. The budget would cut NSF by 56% to $3.9 billion, NASA's science arm by 46% to roughly $3.9 billion, and the National Institutes of Health by a third to around $30 billion. The Department of Energy's Office of Science fared better relative to other science agencies, but it still faces a 14% cut to about $7.1 billion.

Congress is unlikely to implement cuts of that magnitude and will develop its own spending proposals, which in the Senate will need bipartisan support to clear the 60-vote threshold for advancing legislation. During Trump's first presidency, Congress rejected proposed cuts to science agencies and in some cases provided substantial funding increases. In his second term, however, Trump has challenged Congress's spending prerogatives on multiple fronts. (See, for example, PHYSICS TODAY, May 2025, page 22.)

NSF estimates that the cuts in the budget request would result in its total number of competitive grant awards dropping from around 9600 to 2300 and the proposal acceptance rate dropping from 26% to 7%. The cuts would also squeeze the agency's facility operation budget to the point that NSF proposes downsizing actions, such as operating only one of two sites of the Laser Interferometer Gravitational-Wave Observatory. —MA

NSF adds DEI, Israel boycott restrictions to grant terms

In May, NSF updated its grant conditions to bar recipients from operating certain diversity, equity, and inclusion (DEI) programs or participating in boycotts of organizations with ties to Israel. The move follows a similar policy issued by the National Institutes of Health in April. The new conditions state that NSF reserves the right to terminate funding if it deems that recipients "operate any program in violation of federal anti-discriminatory laws or engage in a prohibited boycott."

The new conditions target programs that advance "discriminatory equity ideology," defined by an executive order from President Trump as "an ideology that treats individuals as members of preferred or disfavored groups, rather than as individuals, and minimizes agency, merit, and capability in favor of immoral generalizations." The restriction also applies to programs that "advance or promote DEI" but does not define the scope of that phrase. DEI restrictions implemented by other grant-making agencies have been challenged in court. —CZ

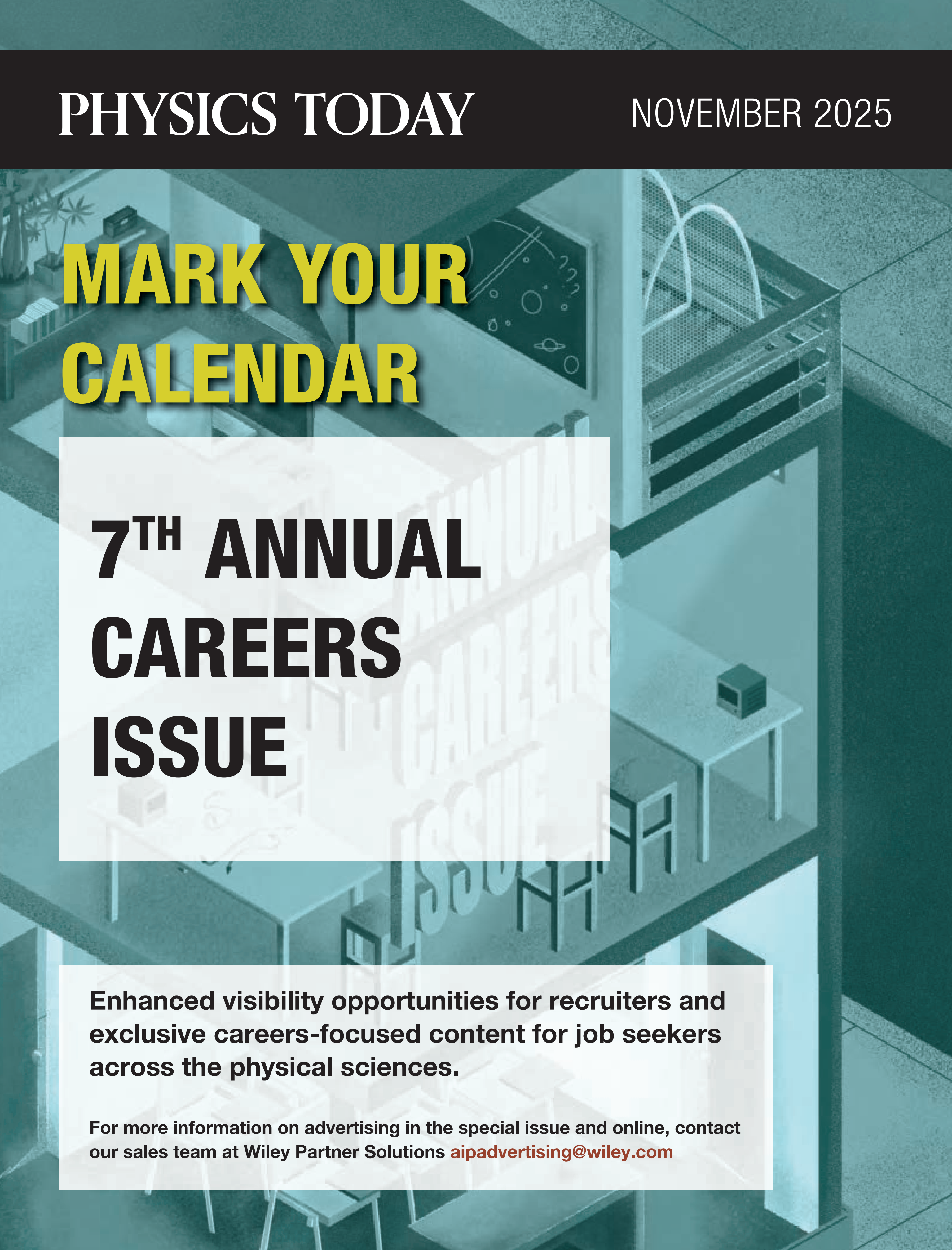
FYI (https://aip.org/fyi), the science policy news service of the American Institute of Physics, focuses on the intersection of policy and the physical sciences.

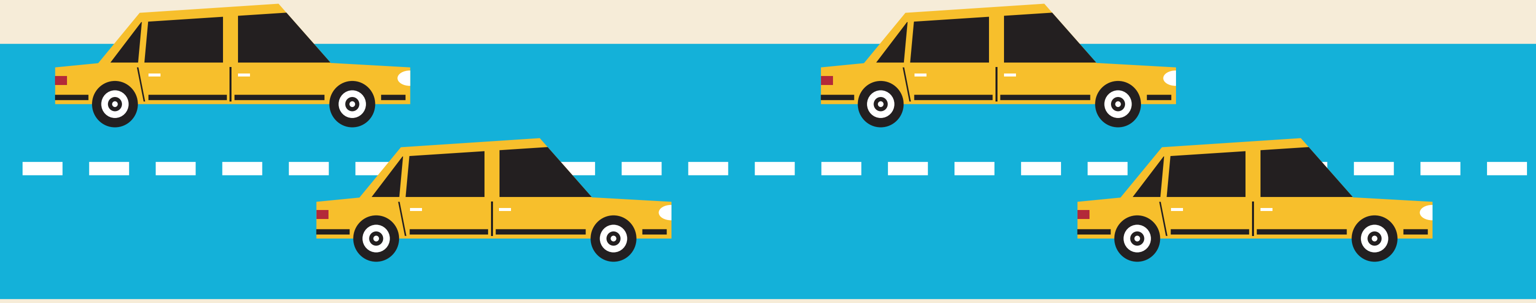
AIP | FYI

AGU, AMS plan climate collection after NCA upheaval

In response to the Trump administration's dismissal of authors working on the latest National Climate Assessment (NCA), the American Geophysical Union and American Meteorological Society plan to solicit submissions for a special collection of recent research on climate change in the US. In April, the administration said that the scope of the report is being reevaluated, and it cut contractor staff that helped coordinate the report writing. About 400 volunteer experts had been working for almost a year on the latest version, which was scheduled to be published near the end of 2027.

The American Geophysical Union and the American Meteorological Society noted in a May press release that the report is congressionally mandated and that the societies' collection of manuscripts "does not replace the NCA but instead creates a mechanism for this important work to continue." They said that their new collection will include more than 29 peer-reviewed journals on climate, and they invited other scholarly publishing organizations to join the effort. (The American Meteorological Society is a member society of the American Institute of Physics, which publishes PHYSICS TODAY.) —CZ

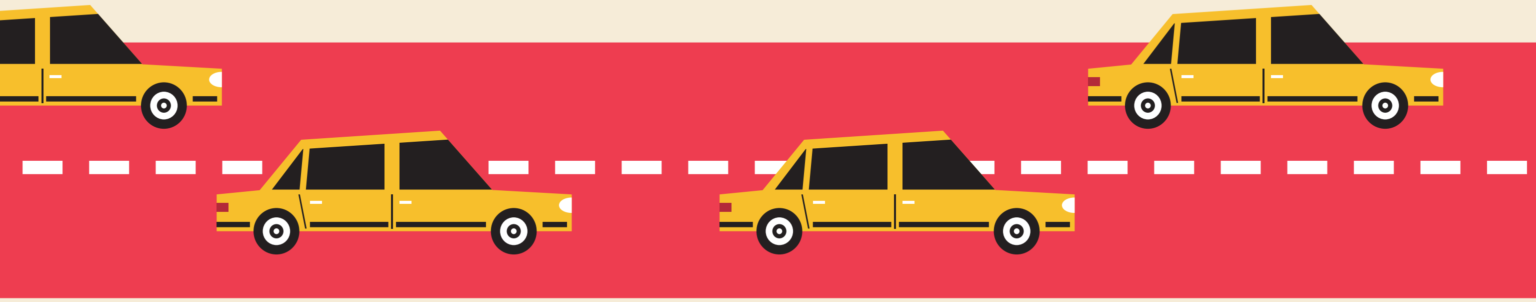

PHYSICS TODAY
NOVEMBER 2025
MARK YOUR CALENDAR
7TH ANNUAL CAREERS ISSUE
Enhanced visibility opportunities for recruiters and exclusive careers-focused content for job seekers across the physical sciences.
For more information on advertising in the special issue and online, contact our sales team at Wiley Partner Solutions aipadvertising@wiley.com



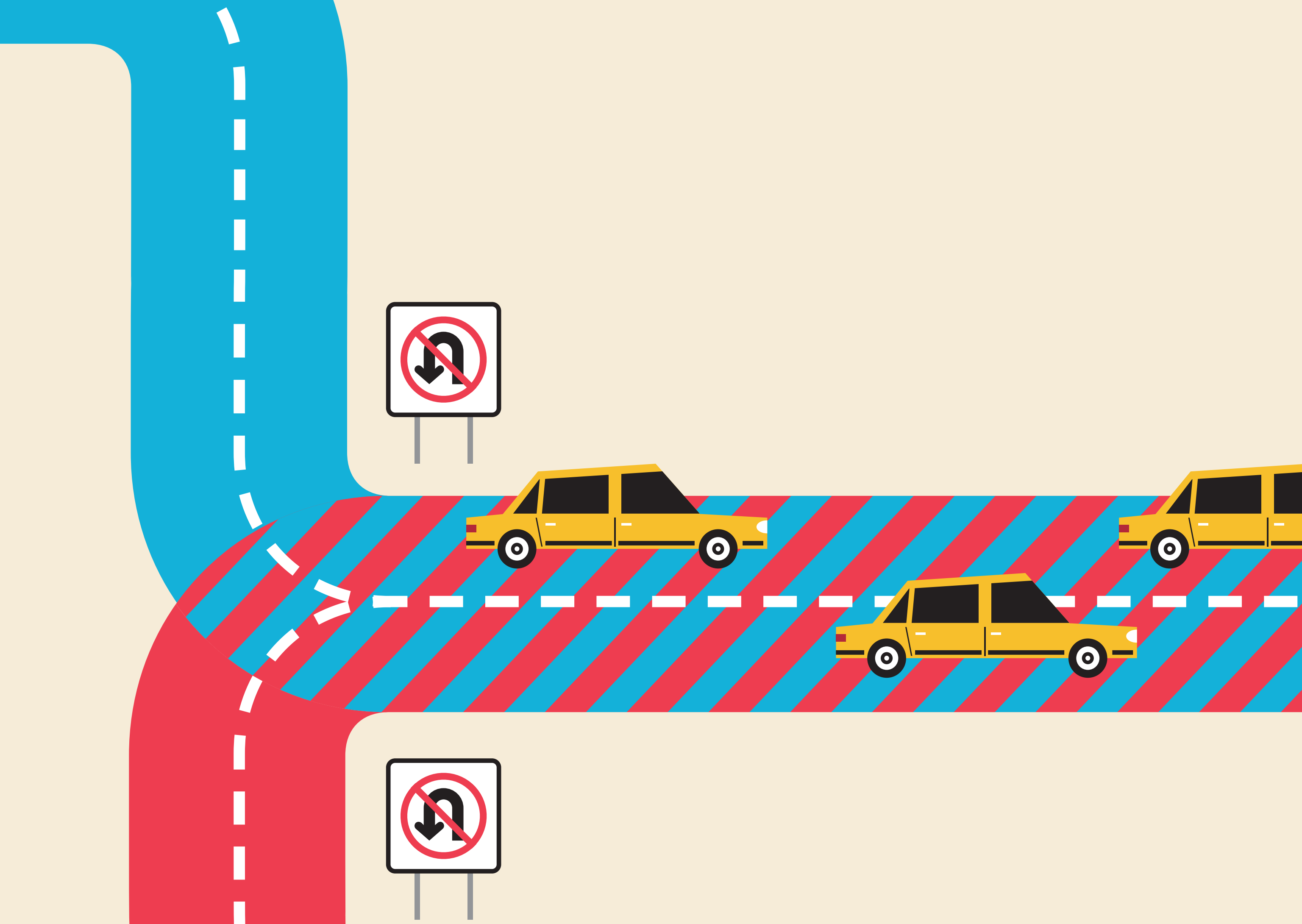
NONINVERTIBLE SYMMETRIES: WHAT'S DONE CANNOT BE UNDONE

Recent research has shown that the traditional notion of symmetry is too limited. A new class of symmetries is bringing surprising insights to quantum systems.

Shu-Heng Shao

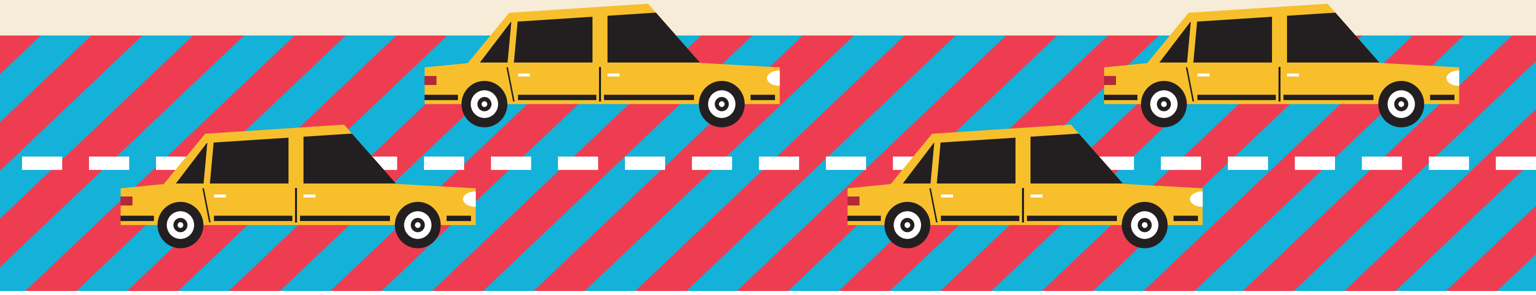


Shu-Heng Shao is an assistant professor at the MIT Center for Theoretical Physics—a Leinweber Institute in Cambridge, Massachusetts. He works on quantum field theory and quantum many-body systems with applications in high-energy physics, condensed-matter physics, and quantum gravity.



Symmetry has long been a foundational concept in theoretical physics and mathematics. It simplifies complex physical problems and reduces the number of unknown variables. You would have a nightmare, for example, solving the Schrödinger equation for the hydrogen atom in Cartesian coordinates *x*, *y*, and *z*. But the problem simplifies dramatically if you use spherical coordinates and leverage the atom's rotational symmetry: The atom looks the same after it's rotated about one of its axes. Symmetry not only helps unify and organize the fundamental forces of nature but also guides the search for new physics.

> The fundamental reason that symmetries can be noninvertible is quantum superposition.



Symmetry transformations are those that leave a system looking and behaving the same. Conventional symmetry transformations are invertible. If we rotate a square by 90°, for example, the transformation can be undone by a −90° rotation. Such intuition is formulated rigorously by Wigner's theorem, which implies that every symmetry transformation in quantum mechanics has an inverse. The mathematical language used to describe conventional symmetry transformations is called group theory, a foundational concept that has shaped modern physics for more than a century (see, for example, the article by Martin Rodriguez-Vega, Maia Vergniory, and Greg Fiete, *Physics Today*, May 2022, page 42).

One way only

In recent years, however, researchers have shown that the traditional notion of symmetry is too limited in quantum field theory and quantum many-body systems. A new class of symmetries—noninvertible—has been identified in various physical systems, including lattice models describing magnetism and quantum field theories of strong interactions between quarks. As the name suggests, noninvertible symmetries are implemented by transformations that do not have inverses—that is, what's done cannot be undone.

The fundamental reason that symmetries can be noninvertible is quantum superposition. In deterministic classical physics, a cat is either alive or dead. In quantum physics, Schrödinger's cat can be both alive and dead simultaneously. The wavefunction describing Schrödinger's cat is a superposition of two individual wavefunctions—one for an alive cat and one for a dead cat. Superposition introduces more possibilities for symmetries in quantum physics: A symmetry transformation can cause the wavefunction of a single cat to become a superposition of two. If the transformation is repeated, the result is a superposition of increasingly many cat wavefunctions, and no inverse transformation reverts to a single cat.

As paradoxical as it may sound, the new symmetries lead to new conservation laws, which serve as novel tools to study strongly coupled physical systems. They also point to alternative physical models and beg for a new mathematical framework to describe symmetries in quantum physics.

Noninvertible symmetry of a magnet

Noninvertible symmetries already exist in physicists' favorite toy model for ferromagnetism: the Ising model in one spatial dimension. The 1D model consists of an array of qubits placed on a circle, as illustrated in figure 1. Each qubit can be spin up $|\uparrow\rangle$, spin down $|\downarrow\rangle$, or any quantum superposition of the up and down states, such as $|\rightarrow\rangle \propto |\uparrow\rangle + |\downarrow\rangle$.

The state $|\uparrow\uparrow\cdots\uparrow\rangle$, in which every spin is pointing up, corresponds to a magnet whose north pole is pointing up. Similarly, the state $|\downarrow\downarrow\cdots\downarrow\rangle$ corresponds to a magnet whose south pole is pointing up. On the other hand, the state

$$|\rightarrow\rightarrow\cdots\rightarrow\rangle \propto |\uparrow\uparrow\cdots\uparrow\rangle + |\downarrow\uparrow\cdots\uparrow\rangle + |\uparrow\downarrow\cdots\uparrow\rangle + \cdots + |\downarrow\downarrow\cdots\downarrow\rangle \quad (1)$$

represents a superposition of all possible spin configurations. Since that is a state with no preference for spin up or spin down, magnetization is lost. The transition from $|\uparrow\uparrow\cdots\uparrow\rangle$ to $|\rightarrow\rightarrow\cdots\rightarrow\rangle$ models the process of heating up a magnet: As the temperature reaches a critical value, the magnet loses its magnetization.

What are the symmetries in the toy model for a magnet? Because the north and south poles are on the same footing, an ordinary symmetry can transform one to the other. It flips all the spins from up to down and vice versa: $|\uparrow\uparrow\cdots\uparrow\rangle \to |\downarrow\downarrow\cdots\downarrow\rangle \to |\uparrow\uparrow\cdots\uparrow\rangle$. That is an invertible symmetry—apply it twice, and we return to the starting point. The demagnetized state $|\rightarrow\rightarrow\cdots\rightarrow\rangle$ is symmetric under the spin-flip symmetry because there is no notion of north versus south.

At the critical temperature, an additional symmetry emerges whose effect is $|\uparrow\uparrow\cdots\uparrow\rangle \to |\rightarrow\rightarrow\cdots\rightarrow\rangle$ and $|\downarrow\downarrow\cdots\downarrow\rangle \to |\rightarrow\rightarrow\cdots\rightarrow\rangle$. The additional symmetry transformation acts identically on the up and down states, as illustrated in figure 2. Whether the output state was initially in the up or down state before the transformation isn't knowable. Relatedly, if we apply the transformation a second time, we find a superposition of up and down states: $|\rightarrow\rightarrow\cdots\rightarrow\rangle \to 1/\sqrt{2}\,(|\uparrow\uparrow\cdots\uparrow\rangle + |\downarrow\downarrow\cdots\downarrow\rangle)$. The symmetry transformation cannot be inverted and thus it is a noninvertible symmetry.

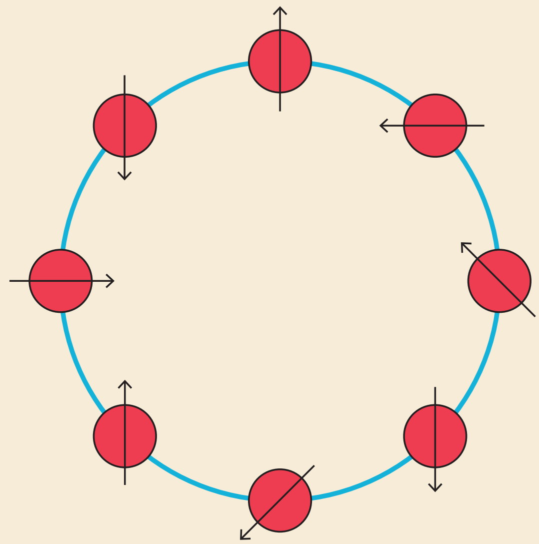


FIGURE 1. THE ISING MODEL consists of an array of qubits in one spatial dimension. Each one can be in a spin-up state $|\uparrow\rangle$, a spin-down state $|\downarrow\rangle$, or a superposition of the two. The model is an archetypal system that explores the differences between ordinary, invertible symmetries and noninvertible symmetries, which, once applied, cannot be undone. (Illustration by Three Ring Studio.)

The technical details

We now examine more carefully the noninvertible symmetry in the Ising model, which has L qubits, labeled by $j = 1, 2, \ldots, L$, arranged on a 1D closed, periodic ring. (It has a counterpart in two dimensions.[1]) On each qubit, a quantum operator, denoted as $\mathbf{Z}$, can be applied to measure the spin: $\mathbf{Z}|\uparrow\rangle = +|\uparrow\rangle$, $\mathbf{Z}|\downarrow\rangle = -|\downarrow\rangle$. Alternatively, another quantum operator, denoted as $\mathbf{X}$, can be applied to flip the spin, where $\mathbf{X}|\uparrow\rangle = |\downarrow\rangle$, $\mathbf{X}|\downarrow\rangle = |\uparrow\rangle$.

If we represent a qubit's spin-up and spin-down states as two column vectors, $\binom{1}{0}$ and $\binom{0}{1}$, then the operators become the Pauli matrices $\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. When we have multiple qubits, we can similarly define $\mathbf{Z}_j$ and $\mathbf{X}_j$ as operators that measure or flip the jth qubit while leaving the others unchanged.

In quantum mechanics, the time evolution of a system is governed by an operator called the Hamiltonian, which features in the Schrödinger equation. At the critical temperature, the Hamiltonian for the Ising model takes the form

$$\mathbf{H} = -\sum_{j=1}^{L} \mathbf{X}_j - \sum_{j=1}^{L} \mathbf{Z}_j \mathbf{Z}_{j+1}\,. \tag{2}$$

The first term models a transverse magnetic field, and the second term models the coupling between the spins of neighboring qubits.

What are the symmetries of the critical Ising model? A necessary condition for a symmetry is that it must lead to a transformation that leaves the Hamiltonian invariant. The Hamiltonian is invariant under the transformation $\mathbf{X}_j \to \mathbf{X}_j$ and $\mathbf{Z}_j \to -\mathbf{Z}_j$.

That transformation is the spin-flip symmetry implemented by the operator $\mathbf{V} = \mathbf{X}_1\mathbf{X}_2 \ldots \mathbf{X}_L$. It commutes with the Hamiltonian, which means that it does not change over time. In other words, it's a conserved quantity.

Is there an additional symmetry in the Ising model? Another transformation that leaves the Hamiltonian invariant is

$$\mathbf{X}_j \to \mathbf{Z}_j\mathbf{Z}_{j+1}\,,\ \mathbf{Z}_j\mathbf{Z}_{j+1} \to \mathbf{X}_{j+1}\,, \tag{3}$$

which is known as the Kramers–Wannier transformation.[2] What exactly do the arrows in equation 3 mean? Even though the literature has commonly suggested that the Kramers–Wannier transformation is invertible, it's not. To see why, let us assume that the transformation is implemented by conjugating an operator by an invertible operator $\mathbf{U}$. The Kramers–Wannier transformation would thus be written as

$$\mathbf{U}\mathbf{X}_j\mathbf{U}^{-1} \stackrel{?}{=} \mathbf{Z}_j\mathbf{Z}_{j+1}\,,\ \mathbf{U}\mathbf{Z}_j\mathbf{Z}_{j+1}\mathbf{U}^{-1} \stackrel{?}{=} \mathbf{X}_{j+1}\,. \tag{4}$$

Let us apply the invertible transformation on the spin-flip operator $\mathbf{V}$ of the Ising model: $\mathbf{U}\mathbf{V}\mathbf{U}^{-1} = \mathbf{U}(\mathbf{X}_1\mathbf{X}_2 \ldots \mathbf{X}_L)\mathbf{U}^{-1} = (\mathbf{Z}_1\mathbf{Z}_2)(\mathbf{Z}_2\mathbf{Z}_3)\ldots(\mathbf{Z}_L\mathbf{Z}_1) = 1$, where in the last step, we have used $\mathbf{Z}_j^2 = 1$. When we multiply by $\mathbf{U}^{-1}$ from the left and $\mathbf{U}$ from the right, the spin-flip operator $\mathbf{V}$ becomes a trivial operator, which is a contradiction.

The Kramers–Wannier transformation, therefore, cannot be implemented by an invertible operator, as Wigner's theorem suggests. Rather, the meaning behind the arrows in equation 3 is answered by the following operator $\mathbf{D}$:[3]

$$\mathbf{D} = \underbrace{e^{-\frac{2\pi i L}{8}} \left(\prod_{j=1}^{L-1} e^{\frac{i\pi \mathbf{X}_j}{4}} e^{\frac{i\pi \mathbf{Z}_j\mathbf{Z}_{j+1}}{4}} \right) e^{\frac{i\pi \mathbf{X}_L}{4}}}_{\text{invertible but not conserved}} \times \underbrace{\frac{1 + \prod_{j=1}^{L} \mathbf{X}_j}{2}}_{\text{conserved but noninvertible}}\,. \tag{5}$$

It is rather complicated, but the only thing we need to know is that $\mathbf{D}$ is a product of an invertible but not conserved operator and a conserved but noninvertible operator. Because of the second factor, $\mathbf{D}$ is a noninvertible matrix that has some zero eigenvalues. The noninvertible operator implements the Kramers–Wannier transformation in the following precise sense:

$$\mathbf{D}\mathbf{X}_j = \mathbf{Z}_j\mathbf{Z}_{j+1}\mathbf{D}\,,\ \mathbf{D}\mathbf{Z}_j\mathbf{Z}_{j+1} = \mathbf{X}_{j+1}\mathbf{D}\,. \tag{6}$$

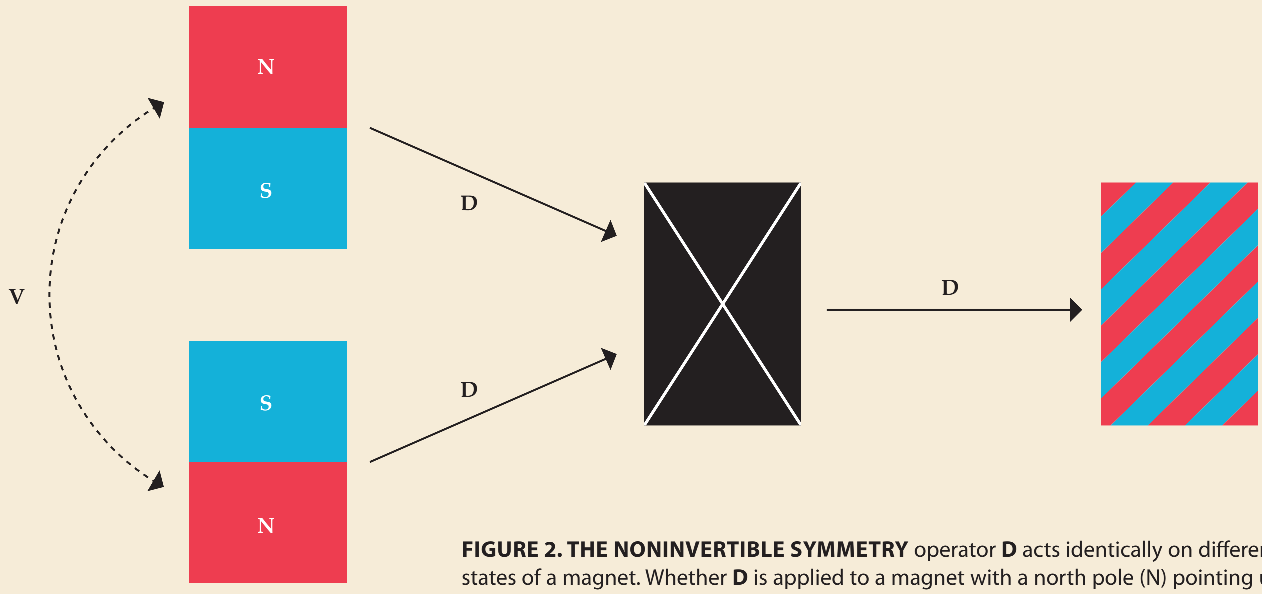


FIGURE 2. THE NONINVERTIBLE SYMMETRY operator **D** acts identically on different states of a magnet. Whether **D** is applied to a magnet with a north pole (N) pointing up or a south pole (S) pointing up, the result is the same: a demagnetized output state, shown in black. When **D** is applied a second time, the magnet enters a quantum superposition of the two initial states. (Illustration by Three Ring Studio.)

Since $\mathbf{D}^{-1}$ does not exist, however, the equation cannot be written in the form of equation 4, and thus no contradiction exists. The operator commutes with the Hamiltonian, and it is therefore a conserved quantity that does not change over time. It is a noninvertible symmetry.

The square of an ordinary symmetry is another symmetry: If a 90° rotation is applied twice, the result is a 180° rotation. But what about for the noninvertible symmetry? From equation 4, we see that applying **D** twice moves $\mathbf{X}_j$ forward to site $j + 1$. It appears that the noninvertible symmetry behaves like a lattice translation by half a site. That's inaccurate, however: Nothing exists between two lattice sites!

In fact, the operator **D**, the spin-flip operator **V**, and the one-site lattice translation operator **T** obey the algebraic relation $\mathbf{D} \times \mathbf{D} = \mathbf{T} \times (1 + \mathbf{V})/2$.

The action of **D** on a state, therefore, is conditioned on the response of that state to the spin-flip symmetry. The square of **D** corresponds to a one-site lattice translation **T** on states that are symmetric under **V**, such as the state in equation 1. But it is zero on states that flip signs under **V**. That algebra, which was recently derived,[3] does not fit in the mathematical framework of group theory and goes beyond the paradigm of Wigner's theorem. The new, noninvertible symmetry brings fresh insights into physical systems more generally.

So what is it good for?

One important task in condensed-matter physics is to characterize the phase diagram of a physical system. A familiar example is water at atmospheric pressure, which has gas, liquid, and solid phases. The task is often challenging because of strong interactions among the microscopic particles and atoms. Symmetry provides one of the few powerful analytic tools available to study such strongly coupled systems.

In particular, the noninvertible symmetries of systems that are invariant under the Kramers–Wannier transformation bring fresh insights into quantum systems. In the critical Ising model and a large class of related models, for example, magnetization and demagnetization coexist, which implies nontrivial entanglement properties. A formalization of that intuition[4-7] shows that the presence of a noninvertible symmetry forbids a featureless phase in which there is no entanglement. Moreover, the symmetry constrains the number of ground states with the lowest energy. Such a constraint would not have been possible if the Kramers–Wannier transformation were mistaken for an ordinary, invertible symmetry.

For many years, the discussion of noninvertible symmetries was confined to toy models in one spatial dimension, such as the Ising model for magnetization. A pair of papers[8,9] from a few years ago, however, led to many developments. Inspired by earlier work,[10] they introduced a construction of noninvertible symmetries applicable in three or more spatial dimensions. The key to the construction was the connection to another type of novel symmetry, known as the higher-form symmetry, that acts on extended objects such as strings.[11]

The ideas led to a rapid discovery of new symmetries across various physical systems, including quantum electrodynamics.[12,13] The symmetries provide tantalizing insights into other topics too, including particle physics; point to mistakes in the literature on scattering amplitudes;[14] and consolidate conjectures in quantum gravity and string theory. Beyond high-energy physics, noninvertible symmetries have also been applied to lattice models in condensed-matter theory and quantum information, which has led to the discovery of novel topological phases of quantum matter and constraints on phase diagrams. The new symmetries have emerged as a unifying language that brings together researchers of high-energy physics, condensed matter, and quantum information.

The idea of the noninvertible symmetry has emerged from interdisciplinary developments in physics and math. Because it goes beyond the framework of group theory, it calls for a new mathematical language. In some cases, the correct language has been identified as category theory, a generalization of group theory. Such advancements are fostering vibrant collaborations between mathematicians and physicists and mark a new chapter in the alignment of the two fields.

Throughout history, symmetries have contributed to major breakthroughs in physics. The symmetry discovered in 1941 by Kramers and Wannier[2] is now understood as a special example of noninvertible symmetries, and it predicted the critical temperature of the Ising model. The result encouraged Lars Onsager in 1944 to find the exact solution of the Ising model.[15] In recent years, the discovery and application of new noninvertible symmetries has led to a wave of progress across various areas of physics, and many more promising breakthroughs are on the horizon.

REFERENCES

1. D. Aasen, R. S. K. Mong, P. Fendley, *J. Phys. A: Math. Theor.* **49**, 354001 (2016).
2. H. A. Kramers, G. H. Wannier, *Phys. Rev.* **60**, 252 (1941).
3. N. Seiberg, S.-H. Shao, *SciPost Phys.* **16**, 064 (2024).
4. E. O'Brien, P. Fendley, *Phys. Rev. Lett.* **120**, 206403 (2018).
5. C.-M. Chang et al., *J. High Energy Phys.* **2019**, 26 (2019).
6. M. Levin, *Commun. Math. Phys.* **378**, 1081 (2020).
7. N. Seiberg, S. Seifnashri, S.-H. Shao, *SciPost Phys.* **16**, 154 (2024).
8. Y. Choi et al., *Phys. Rev. D* **105**, 125016 (2022).
9. J. Kaidi, K. Ohmori, Y. Zheng, *Phys. Rev. Lett.* **128**, 111601 (2022).
10. M. Koide, Y. Nagoya, S. Yamaguchi, *Prog. Theor. Exp. Phys.* **2022**, 013B03 (2022).
11. D. Gaiotto et al., *J. High Energy Phys.* **2015**, 172 (2015).
12. Y. Choi, H. T. Lam, S.-H. Shao, *Phys. Rev. Lett.* **129**, 161601 (2022).
13. C. Córdova, K. Ohmori, *Phys. Rev. X* **13**, 011034 (2023).
14. C. Copetti, L. Córdova, S. Komatsu, *Phys. Rev. Lett.* **133**, 181601 (2024).
15. L. Onsager, *Phys. Rev.* **65**, 117 (1944). PT

INTERNATIONAL YEAR OF
Quantum Science
and Technology

HIGHLIGHTING

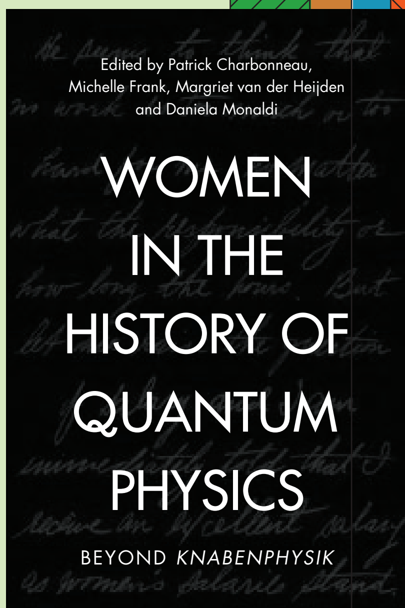
WOMEN IN QUANTUM HISTORY

RYAN DAHN

Ryan Dahn is a senior associate editor at PHYSICS TODAY and a historian of science.

The authors of a new book tell the stories of 16 women who made crucial contributions to quantum physics yet whose names don't usually appear in textbooks.

As modern quantum mechanics was taking shape in the mid 1920s, the field was sometimes referred to in German as *Knabenphysik*—"boys' physics"—because so many of the theorists who were crucial to its development were young men. A new book published as part of the International Year of Quantum Science and Technology pushes back against that male-dominated perspective, which has also tended to dominate historical analyses. Coedited by historians of science Daniela Monaldi and Michelle Frank, physicist-turned-science writer Margriet van der Heijden, and physicist Patrick Charbonneau, *Women in the History of Quantum Physics: Beyond Knabenphysik* presents biographies of 16 oft-overlooked women in the field's history.



The editors did not profile physicists such as Lise Meitner and Maria Goeppert Mayer, who have attracted significant attention from historians and physicists. As the editors explain in the book's introduction, focusing on a few heroic figures perpetuates "a mythology of uniqueness." They instead highlight individuals who are lesser known but nevertheless made important contributions. The following photo essay highlights six of those scientists.

H. JOHANNA VAN LEEUWEN

In 1919, Dutch physicist H. Johanna van Leeuwen (1887–1974) discovered that magnetism in solids cannot solely be explained by classical mechanics and statistical mechanics: It must be a quantum property. Niels Bohr had made the same insight in his 1911 doctoral thesis, but he never published the result in a scientific journal; it was published only in Danish and barely circulated outside Denmark. Van Leeuwen rediscovered what is now called the Bohr–Van Leeuwen theorem in her doctoral research at Leiden University. The theorem, which has applications in plasma physics and other fields, came to the attention of the broader community after Van Leeuwen published an article based on her doctoral thesis in the French *Journal de Physique et le Radium* (Journal of physics and radium) in 1921.

As happened with many women of that era, little trace was left of Van Leeuwen (pictured here in in an undated photo) in the historical record. Chapter authors Van der Heijden and Miriam Blaauboer uncovered several sources that helped them assemble an illustrative synopsis of her career. Van Leeuwen was one of four women to study with Hendrik Lorentz, with whom she remained close until his death in 1928. Unlike many women of her generation, Van Leeuwen remained in the field for her entire career: She was appointed as an assistant at the Technical College of Delft in 1920, a position that required her to supervise laboratory courses for electrical engineering students. In the little spare time she had, Van Leeuwen continued her research into magnetism. In 1947, she was promoted to reader, which meant that she could finally teach her own courses. (Photo courtesy of the Van Leeuwen family.)

LAURA CHALK ROWLES

Laura Chalk Rowles (1904–96) was one of the first women to receive a PhD in physics from McGill University in Montreal, in 1928. Her dissertation investigated the Stark effect—the shifting of the spectral lines of atoms exposed to an external electric field—in the hydrogen atom. In his series of 1926 articles on wave mechanics, Erwin Schrödinger had used quantum theory to predict how the Stark effect would affect the intensities of the Balmer series of spectral lines in hydrogen. As chapter author Daniela Monaldi outlines, Chalk (pictured ca. 1931) used an instrument known as a Lo Surdo tube to measure the intensities of the spectral lines; the work provided the first experimental confirmation of Schrödinger's predictions. She published several articles on the subject in collaboration with her adviser, John Stuart Foster.

Later in his life, Foster regarded his subsequent work on the Stark effect in helium as more important than the hydrogen experiments he had carried out with Chalk. Observers and historians have tended to follow his lead, so her contributions are often overlooked. After spending the 1929–30 academic year at King's College London, Chalk received a teaching position in McGill's agriculture college. But after she married William Rowles, who was also at McGill, she scaled back to working only part time. Five years later, she was let go because of rules that were ostensibly designed to prevent nepotism but typically served to exclude women from the professoriat. (Photo courtesy of Marilyn MacGregor.)

Elizabeth Monroe Boggs (1913–96) received significant press attention for her advocacy work on behalf of people with disabilities. But her prior career in science has long gone overlooked, writes chapter author Charbonneau. Boggs (pictured in 1928) was the only undergraduate to study with famed mathematician Emmy Noether at Bryn Mawr College before Noether's untimely death in 1935. After graduating, Boggs pursued a PhD at the University of Cambridge, where she began studying the application of quantum physics to molecular structure—a pursuit that is now known as quantum chemistry. For her thesis, she used an analog computing device called a differential analyzer to probe the wavefunctions of diatomic molecules.

After finishing her studies, she received a research assistantship at Cornell University, where she met and married chemist Fitzhugh Boggs. As was common in the day, his career took precedence over hers: They moved to Pittsburgh in 1942 when Fitzhugh received a job at Westinghouse. Elizabeth taught at the University of Pittsburgh for a year and then got a job at the Explosives Research Laboratory outside the city, where she ended up contributing to the Manhattan Project by helping to design the explosive lens for implosion bombs like the one ultimately used on Nagasaki. She eventually decided to withdraw from the field and focus on advocacy after the birth in 1945 of a son, David, who had severe developmental delays because of brain damage from an illness. (Photo courtesy of Pamela Murphy.)

ELIZABETH MONROE BOGGS

Katharine Way (1903–95) was the first graduate student of John Wheeler's at the University of North Carolina at Chapel Hill in the late 1930s. As chapter author Stefano Furlan recounts, Way's research during her PhD studies included using the liquid-drop model of the atom, which approximates the nucleus as a droplet of liquid, to examine how nuclei deform when rotating at high speeds. In a 1939 *Physical Review* article, she describes the magnetic moments of heavier nuclei. While carrying out the research, Way (pictured in an undated photo) noticed an anomaly that she brought to Wheeler's attention: The model was unable to account for highly charged nuclei rotating at extremely high speeds. In later recollections, Wheeler regretted that the two didn't further investigate that observation: He noted that, in retrospect, the model's failure in that case was an early indication that nuclei could come apart, just as they do in fission.

During World War II, Way worked on nuclear reactor design at the Metallurgical Laboratory in Chicago; she moved to Oak Ridge Laboratory in 1945. Along with Eugene Wigner, she published a 1948 *Physical Review* article outlining what is now known as the Way–Wigner formula for nuclear decay, which calculates rates of beta decay in fission reactions. She spent much of her postwar career at the National Bureau of Standards (now NIST), where she initiated and led the Nuclear Data Project, a crucial source for information on atomic and nuclear properties that is now part of the National Nuclear Data Center at Brookhaven National Laboratory. Way was also active in efforts to get nuclear scientists to think about the societal ramifications of their work. (Photo courtesy of the AIP Emilio Segrè Visual Archives, Wheeler Collection.)

KATHARINE WAY

SONJA ASHAUER

Although her death from pneumonia at age 25 ended her career practically before it began, Sonja Ashauer (1923–48) was an accomplished physicist and promising talent, chapter authors Barbra Miguele and Ivã Gurgel argue. The daughter of German immigrants to Brazil, Ashauer (pictured ca. 1940) studied at the University of São Paulo with Italian physicist Gleb Wataghin, who likely introduced her to quantum theory. Shortly before the end of World War II, in 1945, she moved to the University of Cambridge, where she became the only woman among Paul Dirac's few graduate students.

In her 1947 thesis, Ashauer worked on one of the most pressing problems of the day in quantum electrodynamics: what was termed the divergence of the electron's self-energy. Because that self-energy—the energy resulting from the electron's interactions with its own electromagnetic field—is inversely proportional to its radius, the value tends to infinity when the particle is modeled as a point charge. Ashauer attacked the problem by working to improve classical electrodynamics in the hope that it might inform the quantum theory. That divergence problem and others were ultimately solved through the renormalization techniques discovered around 1950. (Photo courtesy of the Ashauer family.)

FREDA FRIEDMAN SALZMAN

Freda Friedman Salzman (1927–81) is more often remembered for her work advocating for women in science than for her significant contributions to physics. As an undergraduate, Salzman (pictured in the late 1940s) studied physics with nuclear physicist Melba Phillips at Brooklyn College. In the mid 1950s, in collaboration with her husband, George Salzman, she came up with a numerical method to solve the integral equations of what was known as the Chew–Low model: a description of nuclear interactions developed by Geoffrey Chew—Freda's dissertation adviser at the University of Illinois Urbana-Champaign—and Francis Low. To carry out those calculations, the Salzmans used the ILLIAC I, an early computer. Published in 1957 in *Physical Review*, what was soon termed the Chew-Low-Salzman method helped stimulate work by nuclear and particle physicists, including Stanley Mandelstam, Kenneth Wilson, and Andrzej Kotański in the late 1950s and early 1960s. Chapter author Jens Salomon argues that the method was one of Freda's most important contributions to the field.

Freda and George lived an itinerant academic lifestyle for a period before finding what they believed to be permanent positions at the University of Massachusetts Boston in 1965. Four years later, Freda was fired after the university began to enforce what they claimed to be an anti-nepotism policy. Her termination became a cause célèbre, and after a long campaign, she got her job back in 1972 and received tenure in 1975. The fight to regain her job at the university appears to have motivated Salzman to devote increasing amounts of time to feminist advocacy in the 1970s. (Photo courtesy of Amy Parker.) PT

INTERNATIONAL YEAR OF
Quantum Science
and Technology

Battling Decoherence: The Fault-Tolerant Quantum Computer

Quantum computers have the potential to do certain calculations faster than any foreseeable classical computers, but their success will depend on preserving complex coherent quantum states. Recent discoveries have shown us how to do that.

John Preskill

Editor's note: We continue our celebration of the International Year of Quantum Science and Technology with this introduction to an essential step in building a working quantum computer.

Information carried by a quantum system has notoriously weird properties. Physicists and engineers are now learning how to put that weirdness to work. Quantum computers, which manipulate quantum states rather than classical bits, may someday be able to perform tasks that would be inconceivable with conventional digital technology. (See the article by Charles H. Bennett, Physics Today, October 1995, page 24, and the "Search and Discovery" report in Physics Today, March 1996, page 21.)

Formidable obstacles must be overcome before large-scale quantum computers can become a reality (see the article by Serge Haroche and Jean-Michel Raimond, Physics Today, August 1996, page 51). A particularly daunting difficulty is that quantum computers are highly susceptible to making errors. The magical power of the quantum computer comes from its ability to process coherent quantum states; but such states are very easily damaged by uncontrolled interactions with the environment—a process called decoherence. In response to the challenge posed by decoherence, the new discipline of quantum error correction has arisen at the interface of physics and computer science. We have learned that quantum states can be cleverly encoded so that the debilitating effects of decoherence, if not too severe, can be resisted.

The power of the quantum computer

The indivisible unit of classical information is the bit, which takes one of the two possible values, 0 or 1. Any amount of classical information can be expressed as a sequence of bits. A classical computer executes a series of simple operations (often called "gates"), each of which acts on a single bit or pair of bits. By executing many gates in succession, the computer can evaluate any Boolean function of a set of input bits.

Quantum information, too, can be reduced to elementary units, called quantum bits or qubits. A qubit is a two-level quantum system (like the spin of an electron). A quantum computer executes a series of elementary quantum gates, each of which is a unitary transformation that acts on a single qubit or pair of qubits. By executing many such gates in succession, the quantum computer can apply a complicated unitary transformation to a particular initial state of a set of qubits. Finally, the qubits can be measured; the measurement outcome is the final result of a quantum computation.

A classical computer can faithfully simulate a quantum computer, so that anything the quantum computer could do, the classical computer could also do. Still, there is a sense in which the quantum computer appears to be a more powerful device: Its simulation by the classical computer is very inefficient. The quantum state of even a modest number of qubits (let's say 100) lives in a Hilbert space of unimaginably large dimension: $2^{100} \sim 10^{30}$. To simulate a typical quantum computation, a classical computer would need to work with matrices of exponentially large size, which would take a very long time. In more physical terms, running a classical simulation of a quantum computer is hard because (as exemplified by John Bell's famous inequalities) correlations among quantum bits are qualitatively different from correlations among classical bits. The exponential explosion in the size of Hilbert space as we increase the number of qubits arises because the correlations among qubits are too weird to be expressed easily in classical language.

That simulating a quantum computer with a classical computer takes an unmanageably long time suggested to

John Preskill is a professor of theoretical physics at Caltech.

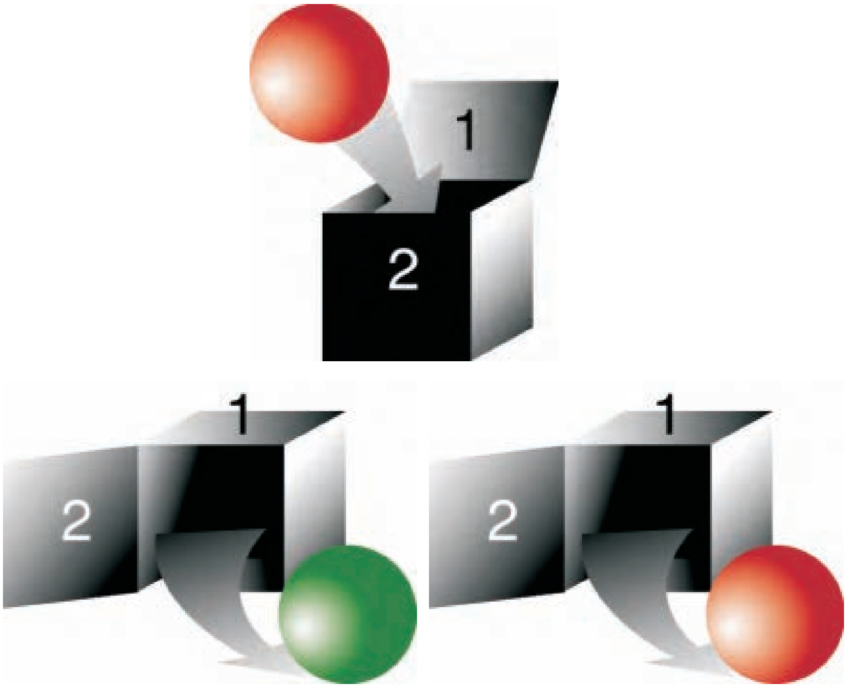


FIGURE 1. DOOR NUMBER 1 or door number 2? To read quantum information reliably, we need to know how it was stored. We can represent an unknown quantum bit (qubit) as a colored ball placed in a box through one of two doors. The doors represent two ways of measuring the qubit (such as the axis along which to measure spin), and the two colors represent the possible outcomes of the measurement. If the ball is placed in the box through door 1, and then it is observed through door 2, the color of the ball that comes out of the box is random.

Richard Feynman[1] that using a quantum computer might enormously speed up finding solutions to certain hard computational problems. David Deutsch,[2] developing the idea further, observed that a quantum computer can invoke a kind of massive parallelism, by operating on a coherent superposition of a vast number of classical states. In fact, a single computation acting on just 300 qubits can achieve the same effect as 2^{300} simultaneous computations acting on classical bits, more than the number of atoms in the visible universe. We could never build a conventional computer with that many processors!

Peter Shor[3] discovered how, in principle, to apply quantum parallelism to the problem of finding the prime factors of a large integer. The difficulty of factoring an integer escalates very rapidly as the number of its digits increases. For example, suppose that we want to find the 65-digit prime factors of a 130-digit composite number. A network of hundreds of powerful workstations, collaborating and communicating over the Internet and running the best algorithms known, might solve the problem in a few months. To factor a 400-digit number, the same network of workstations running the same algorithms would need about 10 billion years (the age of the universe). Even with vast improvements in technology, no one will be factoring 400-digit numbers using conventional computers anytime soon, unless there is an unexpected algorithmic breakthrough.

But now suppose we have a quantum computer that runs just as fast as that network of workstations—that is, it can perform the same number per second of elementary operations on pairs of qubits as the classical computer can perform elementary logic gates on pairs of bits. That quantum computer could factor the 130-digit number in a few seconds, and the 400-digit number in just minutes. Thanks to quantum parallelism, the difficulty scales in a much more reasonable way with the size of the input to the problem. For very large numbers, the advantage enjoyed by the quantum computer is truly stupendous.

The challenge of error correction

If quantum computers would be so marvelous, why don't we just build one? There are technological challenges, to be sure. But are there any obstacles that might be fundamental matters of principle, that would prevent us from ever constructing a quantum computer?

In fact, there is a problem of principle that is potentially very serious: decoherence. Unavoidable interactions with the environment will cause the quantum information stored in a quantum computer to decay, thus inducing errors in the computation. Decoherence occurs very rapidly in complex quantum systems, which is why we never observe macroscopic superpositions (such as a coherent superposition of a live cat and a dead cat). If quantum computers are ever to be capable of solving hard problems, a means must be found to control decoherence and other potential sources of error.

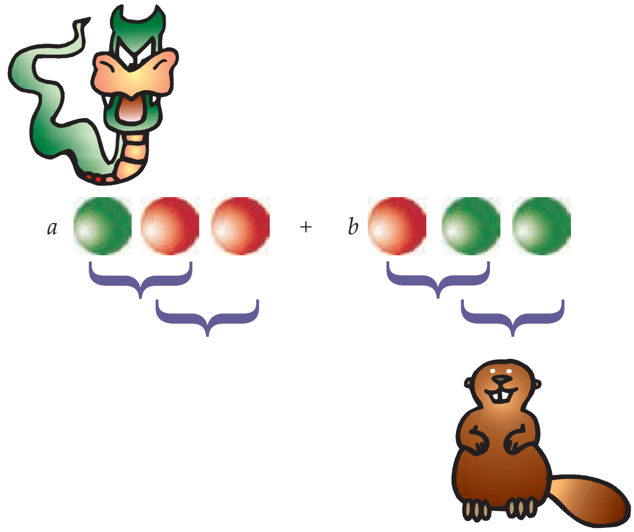


FIGURE 2. ERROR CORRECTION by collective measurement preserves a coherent quantum state. The lurking dragon has flipped one of the three qubits. Measuring two qubits at a time (blue brackets), the busy beaver determines that the first and second qubits are different colors and that the second and third qubits are the same color. He then infers that the first bit has flipped, and repairs the damage.

Errors can be a problem even for classical information. We all have bits that we cherish, while everywhere there are dragons lurking who delight in tampering with our bits. But we have learned some ways to protect classical information from the dragons. If I have a bit with the value 0 that I want to preserve, then I can store two backup copies of the bit. Eventually, a dragon could come along and flip one of my three bits from 0 to 1. But I can employ a busy beaver to check the three bits frequently; when he finds that one has a different value than the others, he flips that bit so that all three match again. That way, as long as the dragon has not had a chance to flip two bits, the error can be corrected and the information will be protected.

We would like to apply the same principle of redundant storage to quantum information, but, because qubits are different from classical bits, there are complications. We might visualize a qubit as a colored ball, either red or green, concealed in a locked box, that can be opened through either of two doors. The doors represent two ways of measuring the qubit, just as we could measure the spin of an electron along either the z or the x axis; the two possible colors represent the possible outcomes of the measurement. If we store a ball in the box through door 1 or door 2 and we later open the same door, we can recover our bit and read it, just as we would read classical information. But if we store the ball through door 1 and then open door 2, what comes out will be completely random (has equal probability of being red or green); the outcome tells us nothing about what we put inside the box (see figure 1). To read quantum information reliably we need to know how it was stored; otherwise we are bound to damage it irrevocably.

The first problem we encounter in the battle against decoherence is that an unknown quantum state cannot be perfectly duplicated;[4] hence we cannot safeguard a quantum computer against errors by storing backup copies of its state. Roughly speaking, the trouble is that to duplicate the information in a quantum box, a copier must open a door to see what is inside. If it just happens to open the same door that was used to store the information, it can make an accurate copy. But if it guesses wrong, it will irrevocably damage the information instead. We can clone a sheep, but not a qubit!

A second problem is that there are more things that can go wrong with quantum information than with classical information. The dragon might open door 1, change the color of the ball, and reclose the box—that would be a *bit-flip* error analogous to the errors that can afflict classical information. Or he might open door 2, change the color, and reclose the box—that would be a *phase* error, for which classical information has no analog. The beaver needs to be able to fix the error without knowing ahead of time whether the dragon is going to use door 1 or door 2.

Third, whereas errors in classical information are discrete, errors in quantum information form a continuum. Rather than simply flipping a bit, the dragon might introduce a more subtle kind of error by performing the bit flip with some (small) probability amplitude ε. The beaver must be able to recover from that kind of small error; otherwise small errors will accumulate over time, eventually building up to become large errors.

Finally, to diagnose whether errors have occurred, the beaver must look at some qubits—and therefore must open some boxes. But quantum measurement necessarily disturbs the state that is being measured, so we worry that the beaver cannot check for errors without introducing further errors.

Quantum error-correcting codes

As recently as four years ago, the difficulties described above seemed highly discouraging. But in 1995, Shor and Andrew Steane discovered[5,6] that the obstacles were illusory—that quantum error correction really is possible. Theirs is one of the most important discoveries about quantum information in recent years, and it can be expected to have far-reaching implications.

To appreciate the insights of Shor and Steane, let's first consider how to defend quantum information against a dragon who performs only bit flips (we'll return to the issue of phase errors shortly). We are to protect the state

$$a|0\rangle + b|1\rangle, \tag{1}$$

a coherent superposition of the red ($|0\rangle$) and green ($|1\rangle$) states of a single qubit, where the complex coefficients a and b are unknown. Were the dragon to attack, the bit flip would transform the state to

$$a|1\rangle + b|0\rangle, \tag{2}$$

and damage would be inflicted unless $a = \pm b$. The beaver's assignment is to diagnose and reverse bit flips, but without disturbing the delicate superposition state, that is without modifying a and b.

Well schooled in classical error correction, the beaver applies the principle of redundant storage by encoding the qubit in a state of three qubits. The red state is encoded as three red qubits, and the green state as three green qubits; that is,

$$|0\rangle \to |\bar{0}\rangle \equiv |000\rangle, \quad |1\rangle \to |\bar{1}\rangle \equiv |111\rangle. \qquad (3)$$

Thus the unknown superposition state becomes

$$a|0\rangle + b|\bar{1}\rangle \to a|\bar{0}\rangle + b|\bar{1}\rangle = a|000\rangle + b|111\rangle. \qquad (4)$$

This redundant state is *not* the same as three identical copies of the original unknown state, which would be

$$(a|0\rangle + b|1\rangle)(a|0\rangle + b|1\rangle)(a|0\rangle + b|1\rangle). \qquad (5)$$

Although it is impossible to copy unknown quantum information, nothing prevents us from building a (unitary) machine that will execute the encoding transformation given as equation 4.

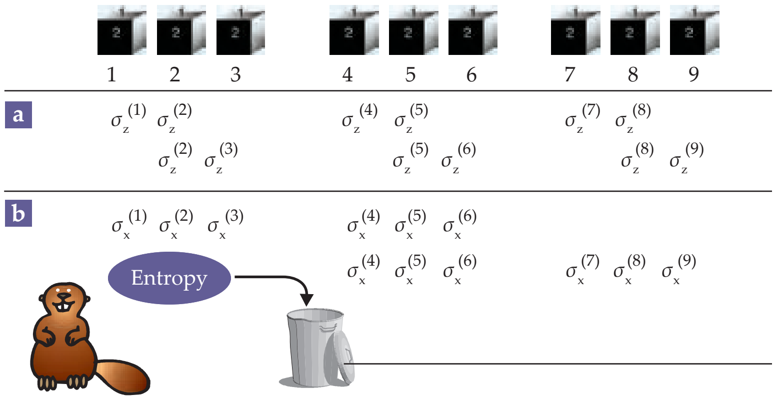


FIGURE 3. A QUANTUM CODE. It is possible to correct both bit-flip and phase-flip errors by encoding one qubit of quantum information in a block of nine qubits. Collective measurements preserve unknown individual qubit states (represented by closed boxes). **(a)** Six two-qubit observables (such as the tensor product of Pauli matrices $\sigma_z^{(1)} \otimes \sigma_z^{(2)}$) are measured to diagnose bit flips. **(b)** Two six-qubit observables (such as the tensor product of Pauli matrices $\sigma_x^{(1)} \otimes \sigma_x^{(2)} \otimes \sigma_x^{(3)} \otimes \sigma_x^{(4)} \otimes \sigma_x^{(5)} \otimes \sigma_x^{(6)}$) are measured to diagnose phase flips. Entropy introduced by errors is extracted in the form of a random measurement record, which can be discarded.

Now suppose that the dragon flips one of the three qubits, let's say the first one, so that the state becomes

$$a|100\rangle + b|011\rangle, \qquad (6)$$

and the beaver is to detect and reverse the damage. His first impulse would be to open the boxes and look to see if one ball was a different color from the others, just as he would to diagnose errors in classical information, but he must resist that temptation. If he were to open door 1 of all three boxes, he would find either $|100\rangle$ (with probability $|a|^2$), or $|011\rangle$ (with probability $|b|^2$); either way, the coherent quantum information (the values of a and b) would be irrevocably lost.

But he is a clever beaver who knows he need not restrict his attention to single-qubit measurements. Instead, he performs collective measurements on two qubits at once (see figure 2). The beaver asks whether the first two qubits have the same color or different colors, without trying to ascertain the color of either one. He finds that the colors are different. Then he asks whether the second and third qubits have the same color or different colors. He finds that the colors are the same. From the two measurement outcomes, the beaver infers that the first qubit has flipped relative to the other two and should be flipped back to repair the damage. In executing this protocol, the beaver has not learned anything about

Box 1. Fault Tolerance and Topology

Topological ideas arise naturally in the theory of fault tolerance. The topological properties of an object remain invariant when we smoothly deform the object. Similarly, how a fault-tolerant gate acts on encoded information should remain unchanged when we deform the gate by introducing a small amount of noise. In seeking fault-tolerant implementations of quantum logic, we are led to contemplate physical interactions with a topological character.

What comes quickly to mind is the Aharonov–Bohm effect. When an electron is transported around a magnetic flux tube, its wave function acquires a phase that depends only on the winding number of the electron about the solenoid; it is unmodified if the electron's trajectory is slightly deformed. A device that processes quantum information by means of Aharonov–Bohm interactions would be intrinsically fault tolerant; accordingly, we would not need to implement a quantum gate with great precision for it to act as we desire.

Unfortunately, the Aharonov–Bohm effect is abelian, and we need noncommuting gates to build up a complex quantum computation. But it is possible in principle to devise two-dimensional spin systems that exhibit more intricate Aharonov–Bohm phenomena; long-range quantum correlations in the ground state of such a system can induce topological interactions among the localized quasiparticle excitations.[12] In a suitable spin system, the Aharonov–Bohm interactions are adequate for executing interesting computations like the quantum factoring algorithm.

Such an implementation of quantum computation seems futuristic from the perspective of current technology, but it is conceptually important. If we could perform quantum logic by means of topological interactions, then we would be able to give the beaver a rest! We could protect encoded information not by vigilantly checking for errors and reversing them, but rather by weaving fault tolerance into the design of our hardware.

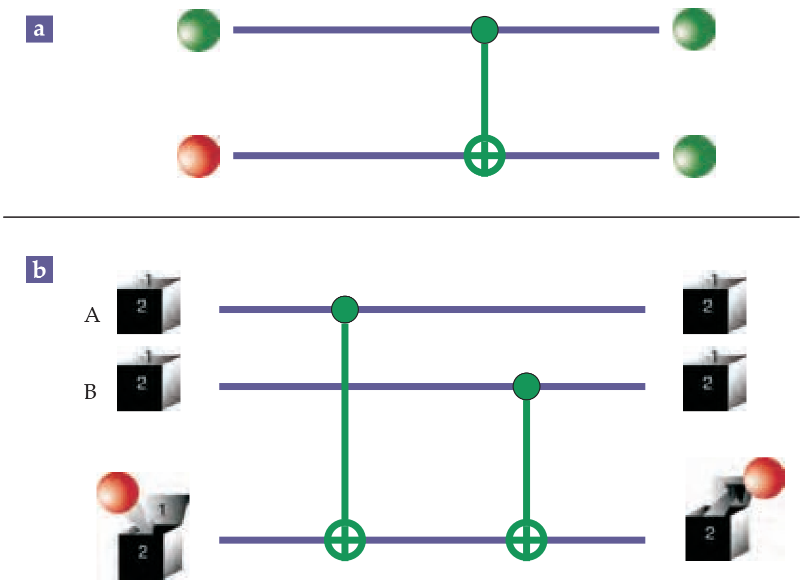


FIGURE 4. QUANTUM LOGIC of a collective measurement. **(a)** A controlled-NOT gate flips the target qubit if the control qubit (on top) is green. Otherwise, it acts trivially. **(b)** A collective observable of two data qubits (marked A and B) is measured by preparing an ancilla qubit, executing two controlled-NOT gates, and then measuring the ancilla.

the encoded state (the values of a and b), hence the recovery procedure itself has inflicted no damage.

The beaver won that round, but now the dragon tries a more subtle approach. Rather than flipping the first qubit, he rotates it only slightly, so that the three-qubit state becomes

$$a|000\rangle + b|111\rangle \rightarrow a|000\rangle + b|111\rangle + \varepsilon(a|000\rangle + b|111\rangle) + O(\varepsilon^2), \quad (7)$$

where $|\varepsilon| \ll 1$. What should the beaver do now? In fact, he can do the same thing as before. If he performs a collective measurement on the first two qubits, then most of the time (with probability $1 - |\varepsilon|^2$), the measurement will project the damaged state (equation 7) back to the completely undamaged state (equation 4). Only much more rarely (with probability $|\varepsilon|^2$) will the measurement project onto the state given as equation 6 with a bit-flip error. But then the measurement outcome tells the beaver what action to take to repair the damage, just as in the previous case.

Of the four difficulties for quantum error correction cited above, then, we have already seen how three can be overcome. We can encode a quantum state redundantly without violating the no-cloning principle. We can perform collective measurements that let us acquire information about the nature of the errors without revealing anything about the state, and so without damaging the state. We can control the accumulation of small errors by repeatedly making measurements that either reverse the damage or introduce large errors that we know how to correct. It remains only to resolve one more issue: the problem of phase errors.

Fixing phases

The code we have devised so far provides no protection against a dragon who flips the relative phase of $|0\rangle$ and $|1\rangle$. If such a

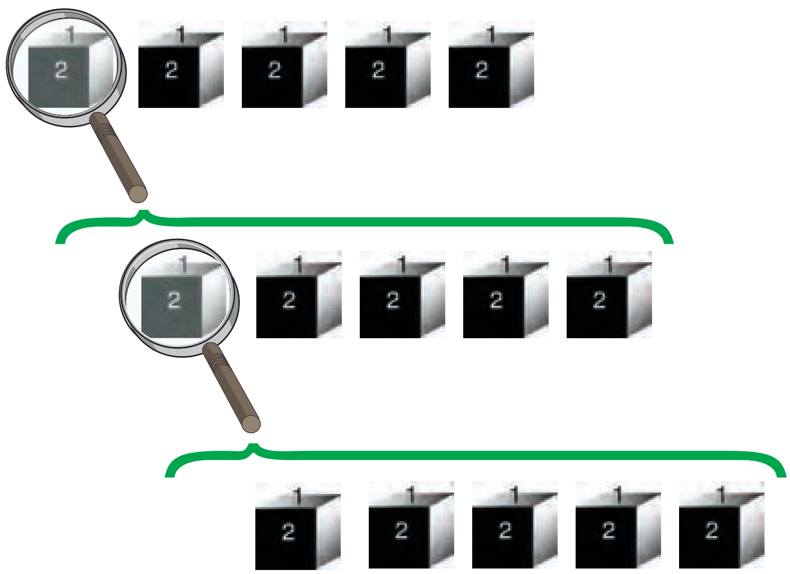


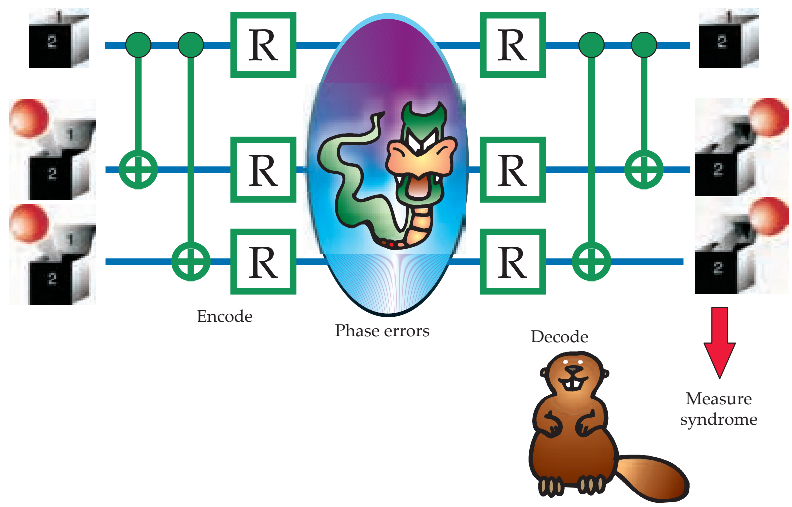
FIGURE 5. CODES WITHIN CODES. A single logical qubit is encoded in a block of five qubits. Each of the five qubits in that block, when inspected at higher resolution, is itself really a block of five qubits. And so on.

Box 2. Experimental Quantum Error Correction

The first experimental demonstrations of quantum error correction, using the methods of nuclear magnetic resonance (NMR), were reported in the past year. In those experiments, qubits were carried by nuclear spins that were manipulated by radiofrequency pulses, and quantum coding was used to protect a spin from dephasing. In an experiment by a group from Los Alamos National Laboratory and MIT,[13] (schematically illustrated in the figure), two ancilla spins were provided, and the qubit to be protected was encoded in correlations among the three by means of a simple quantum circuit. The three spins were exposed to the dephasing dragon for a while, and then the qubit was decoded. The ancilla spins were measured to reveal whether a phase error had been sustained; if it had, the damage could be repaired.

In an experiment conducted by a group from IBM/Almaden and Stanford University,[14] a two-qubit code that could detect a phase error in either qubit was used, and the output was rejected when an error was detected. In the cases in which no error was detected, an improvement in fidelity could be verified.

Quantum error correction demonstrations that exploit the tools of quantum optics and atom trapping should be possible in the near future.[15]



PROTECTING A NUCLEAR SPIN from phase errors. First, some controlled-NOTs and some single-qubit quantum gates are executed to encode the spin to be protected (top left) in correlations with the two ancilla spins (shown below it). Then the three spins, now in an entangled state, are subjected to weak dephasing. Finally, the spins are decoded, and two are measured to extract a syndrome that diagnoses whether a phase error has occurred.

dragon attacks any one of our three qubits, then our encoded state $a|\bar{0}\rangle + b|\bar{1}\rangle$ is transformed to $a|\bar{0}\rangle - b|\bar{1}\rangle$, and the encoded quantum information is damaged if a and b are both nonzero. But the method we developed to conquer the bit-flip errors can be extended to deal with phase errors as well—just as we protected against bit-flip errors by encoding bits redundantly, we can protect against phase-flip errors by encoding phases redundantly.

Following Shor,[5] we may encode a single qubit using a block of nine qubits (see figure 3), according to

$$|0\rangle \to |\bar{0}\rangle \equiv \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle),$$

$$|1\rangle \to |\bar{1}\rangle \equiv \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \quad (8)$$

Both $|\bar{0}\rangle$ and $|\bar{1}\rangle$ consist of three clusters of three qubits each, with each cluster prepared in the same quantum state. Each of the clusters has triple-bit redundancy, so we can correct a single bit flip in any cluster by the method already discussed above.

Now suppose that a phase flip occurs in one of the clusters. The error changes the relative sign of $|000\rangle$ and $|111\rangle$ in that cluster so that

$$|000\rangle + |111\rangle \to |000\rangle - |111\rangle,$$
$$|000\rangle - |111\rangle \to |000\rangle + |111\rangle. \quad (9)$$

The relative phase of the damaged cluster will now differ from the phases of the other two clusters. Thus, we can identify the damaged cluster, not by *measuring* the relative phase in each cluster (which would disturb the encoded information) but by *comparing* the phases of pairs of clusters—a six-qubit collective measurement. The measurement outcomes allow us to infer which cluster has a sign different from the others, and we may then apply a unitary phase transformation to one of the qubits in that cluster to reverse the sign and correct the error.

Error recovery will fail if there are two bit-flip errors in a single cluster (which would induce a phase error in the encoded data) or if phase errors occur in two clusters (which would induce a bit-flip error in the encoded data). But if the qubits interact only weakly with the environment and with one another, a double error will be relatively unlikely. Loosely speaking, if each qubit decoheres with a probability p and the decohering qubits are not strongly correlated, then the encoded information will decohere with a probability of order p^2. For p sufficiently small, coding will improve the reliability of the quantum information.

The nine-qubit code is conceptually simple, but it is not the most efficient quantum code that can protect against an arbitrary error afflicting any one of the qubits in the code block. It turns out that a five-qubit code can be devised to accomplish the same thing.[7] More sophisticated codes can be constructed that can protect against many damaged qubits in the code block.[8]

Collective measurement and fault tolerance

Collective measurements, which can diagnose errors without damaging the coherence of the data, are crucial to quantum error correction. Let's consider more closely how collective measurements can be carried out. The beaver would like to learn, for example, whether boxes *A* and *B* (both opened through door 1) contain balls of the same color or different color, but he doesn't want to find out the color of either ball.

To measure such collective observables, he will need a rudimentary quantum computer that can perform quantum logic gates in which two qubits come together and interact (see figure 4). A two-qubit gate that is particularly useful for this purpose is the controlled-NOT gate that acts according to this rule: If the first (control) qubit is $|0\rangle$, then the gate acts trivially, but if the first qubit is $|1\rangle$, the gate flips the value of the second (target) qubit.

When the beaver wants to measure the collective observable, he first prepares a third ("ancilla") qubit in the red state $|0\rangle$. Then a quantum circuit is executed in which two successive controlled-NOT gates are performed, each with the ancilla as the target and with the successive qubits *A* and *B* as the controls. If qubits *A* and *B* have the same color, the color of the ancilla qubit is flipped either zero times or twice, so it is still red when measured; but if qubits *A* and *B* have different colors, there is only one flip, and the ancilla becomes green. Measuring the ancilla reveals only the collective property, not the colors of the two individual qubits.

The ancilla is an essential part of the quantum error correction procedure, because it serves as a repository for the entropy that is introduced into the code block by the errors—it "heats" as the protected quantum system "cools." To protect quantum information for a long time, we need a continual supply of fresh ancilla qubits. Alternatively, if the ancilla is to be recycled, it must be erased. The erasure is a dissipative process; that is why quantum (or classical) error correction requires the expenditure of power.

Since our quantum computer will not be flawless, errors might occur during the collective measurement. Therefore, we must be careful to design a protocol for error recovery that is fault tolerant, one that will still work effectively even if it is not executed perfectly. Indeed, fault-tolerant protocols can be constructed both for error correction and for executing quantum gates that process the encoded information.[9] Box 1 on page 45 describes a topological approach to fault tolerance.

If we wish to perform a long quantum computation reliably, we will need to use codes that can protect against many errors. One family of such codes can be envisioned as follows[10] (see figure 5): Suppose that we encode a single qubit in a block of five qubits. But each of those five qubits, when inspected more closely, is itself really another block of five, encoded as before. And so on. Such an intricate code requires substantial storage space, but in return we achieve high reliability. For an error to occur in the encoded qubit at the highest level, two qubits in the block of five would need to fail. And for either of those to fail, two would need to fail at the next level down. And so on. As we add more levels to the code, the probability of an error in the encoded qubit drops sharply.

Because of the overhead associated with processing encoded information, if our quantum hardware is highly inaccurate, then coding alone may not improve the performance of a quantum computer. But when the hardware becomes reliable enough, an encoded block will be more resistant to error than a raw qubit. Then adding another level to the code will improve the accuracy further. By using a sufficiently complex code, we can make the error rate in the encoded data as small as we please.[11]

In principle, then, an arbitrarily long quantum computation can be performed reliably, provided that the average probability of error per elementary quantum gate is less than a certain critical value, the accuracy threshold. The numerical value of the accuracy threshold depends on the model of decoherence that we adopt, and on other characteristics of our hardware. If we assume that the quantum hardware is highly parallelizable (so that we can execute many quantum gates in a single time step), and that the qubits decohere more or less independently, then an error probability per gate of 10^{-4} can be shown to be acceptable. (Roughly speaking, this error probability can be interpreted as the ratio of the time required to execute an elementary gate to the decoherence time of a single raw qubit.) Of course, to perform a longer computation, more redundancy will be needed for adequate reliability. But the required block size of the code grows at a modest rate with the length of the computation, as a power of a logarithm of the number of gates to be executed.[11]

Outlook

We may now claim to understand, in principle, how to fight off the destructive effects of decoherence. Though we may never see a real cat in a superposition of a dead state and a live state, someday we may be able to prepare an encoded cat that is half dead and half alive, and to maintain that macroscopic superposition for as long as we please.

At present, though, quantum information technology remains in its pioneering stage. It is currently possible to do experiments involving a few qubits and a few quantum gates (box 2 on page 47). For a quantum computer to compete with a state-of-the-art classical computer, we will need machines with hundreds or thousands of qubits capable of performing millions or billions of operations. The technology clearly has far to go before quantum computers can assume their rightful place as the world's fastest machines. But now that we know how to protect quantum information from errors, there are no known insurmountable obstacles blocking the path. Quantum computers of the 21st century may well unleash the vast computational power woven into the fundamental laws of physics.

Apart from enabling a new technology, the discovery of fault-tolerant methods for quantum error correction and quantum computation may have deep implications for the future of physics. Efficient quantum algorithms (such as Shor's factoring algorithm) demonstrate that quantum systems of modest size can behave in ways that classical systems could never imitate. What else might coherent quantum systems be capable of? In what ways will they surprise, baffle, and delight us? Armed with new tools for maintaining and controlling intricate quantum states, physicists of the next century will seek the answers.

REFERENCES

1. R. P. Feynman, *Int. J. Theor. Phys.* **21**, 467 (1982).
2. D. Deutsch, *Proc. Roy. Soc. London, Ser. A* **400**, 96 (1985).
3. P. Shor, in *Proc. of the 35th Annual Symp. on Foundations of Computer Science*, Los Alamitos, Calif., IEEE Press (1994), p. 124.
4. D. Dieks, *Phys. Lett. A* **92**, 271 (1982). W. K. Wootters, W. H. Zurek, *Nature* **299**, 802 (1982).
5. P. Shor, *Phys. Rev. A* **52**, 2493 (1995).
6. A. M. Steane, *Phys. Rev. Lett.* **77**, 793 (1996).
7. R. Laflamme, C. Miquel, J. P. Paz, W. Zurek, *Phys. Rev. Lett.* **77**, 198 (1996); C. Bennett, D. DiVincenzo, J. Smolin, W. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
8. A. R. Calderbank, P. W. Shor, *Phys. Rev. A* **54**, 1098 (1996); A. M. Steane, *Proc. Roy. Soc. London, Ser. A* **452**, 2551 (1996).
9. P. Shor, in *Proc. of the 37th Annual Symp. on Foundations of Computer Science*, Los Alamitos, Calif., IEEE Press (1996), p. 56.
10. E. Knill, R. Laflamme, https://arxiv.org/abs/quant-ph/9608012.
11. E. Knill, R. Laflamme, W. Zurek, *Proc. Roy. Soc. London, Ser. A* **454**, 365 (1998); A. Yu. Kitaev, *Russ. Math. Surveys* **52**, 1191 (1997); D. Aharonov, M. Ben-Or, in *Proc. of the 29th Annual ACM Symp. on the Theory of Computing*, New York, ACM (1997), p. 176; J. Preskill, *Proc. Roy. Soc. London, Ser. A* **454**, 385 (1998).
12. A. Yu. Kitaev, *Ann. Phys.* **303**, 2 (2003); J. Preskill, in *Introduction to Quantum Computation and Information*, H.-K. Lo, T. Spiller, S. Popescu, eds., p. 213.
13. D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, S. S. Samaroo, *Phys. Rev. Lett.* **81**, 2152 (1998).
14. D. Leung, L. Vandersypen, X. Zhou, M. Sherwood, C. Yannoni, M. Kubinec, I. Chuang, *Phys. Rev. A* **60**, 1924 (1999).
15. C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, D. J. Wineland, *Phys. Rev. Lett.* **75**, 4714 (1995); Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, H. J. Kimble, *Phys. Rev. Lett.* **75**, 4710 (1995). PT

Focus on cryogenics, vacuum equipment, materials, and semiconductors

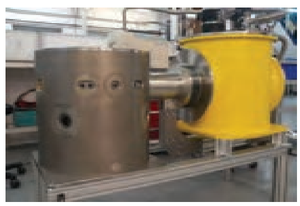
The descriptions of the new products listed in this section are based on information supplied to us by the manufacturers. Physics Today can assume no responsibility for their accuracy. For more information about a particular product, visit the website at the end of its description. Please send all new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

Andreas Mandelis

Electric angle valve

According to Pfeiffer Vacuum+Fab Solutions, its Series E electric angle valve for high-vacuum applications offers superior control, energy efficiency, reliability, and cost savings compared with other available products. Series E is suitable for critical high-vacuum isolation applications ranging from 10^{-3} to 10^{-9} mbar. An efficient 24 V_{DC} electric actuation system and robust elastomer seal ensure precise, reliable operation. Its fail-safe mechanism maintains safety during power losses, and a local mode permits manual operation during maintenance. Those features reduce mechanical stress and ensure consistent performance. The motor-controlled system uses less energy than traditional valves and thus helps reduce operational costs. Designed for easy integration, the right-angle structure enhances system compatibility and user-friendliness. ***Pfeiffer Vacuum Inc****, 24 Trafalgar Sq, Nashua, NH 03063, www.pfeiffer-vacuum.com*

Split cryostat



The Dry ICE Dyad from ICEoxford is a closed-cycle system capable of cooling to 1.4 K while providing excellent optical access and ultralow vibration. The sample unit is held separately from the cold head and the main body of the cryostat; to reduce vibration to less than 20 nm, the two are connected only by a soft thermal link. The sample space environment can be switched between a top-loading exchange gas module and a vacuum module within a couple of hours, which provides the flexibility needed for busy laboratories to address multiple applications. The Dyad features a 5 T split-pair magnet, and five windows with a short working distance provide optical access. It has a base temperature of 1.3 K and provides cooling powers of 50 mW at 1.5 K, 100 mW at 1.6 K, and 200 mW at 1.9 K. According to the company, the specifications of the system make it suitable for studies that may advance the boundaries of quantum optics. ***ICEoxford****, Avenue 4, Station Ln, Witney, Oxfordshire OX28 4BN, UK, www.iceoxford.com*

FTIR system with vacuum ATR accessory

Bruker says that the first product in its new Vertex Neo platform, the Vertex Neo R, represents an advancement in high-end Fourier transform IR (FTIR) research instrumentation for academic and industrial R&D. The user-friendly instrument will enable researchers to explore new frontiers in fields such as catalytic investigations, battery-material development, and semiconductor research. The Vertex Neo R features Bruker's MultiTect detector technology and its first Vacuum ATR (attenuated total reflection) accessory. MultiTect offers broad spectral range detection; the Vacuum ATR builds on Bruker's Platinum ATR accessory, which uses a diamond crystal. It enables ATR measurements with the complete optical path under vacuum and full access to the crystal and the sample at the top of the instrument. That makes venting between measurements unnecessary and allows for complex scientific setups around samples. The ATR accessory enables measurements from the mid- to the far-IR and even the terahertz spectral region, and it is suitable for many types of samples, including solids and volatile liquids. ***Bruker Optics Inc****, 40 Manning Rd, Billerica, MA 01821, www.bruker.com*

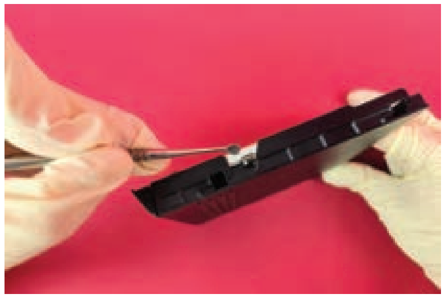
Source measure unit for nanoscale devices

Lake Shore Cryotronics has unveiled the SMU-10, the newest module in its MeasureReady M81-SSM synchronous source measure system. According to the company, it is the first source measure unit optimized for nanoscale devices. Offering very low noise measurements with nanovolt and femtoamp precision, the SMU-10 is suitable for characterizing the ultralow voltage regimes of nanoscale and 2D nanomaterial semiconductors. It provides increased measurement sensitivity by thermal and offset reduction through AC measurement capabilities. Synchronous/simultaneous sourcing and measuring removes sampling misalignment, making the module appropriate for semiconductor I–V, or current–voltage, testing. The SMU-10 also features integrated lock-in detection capabilities. Incorporating six instruments into a unified solution, it provides DC-current, DC-voltage, AC-current, AC-voltage, lock-in, and resistance measurement capabilities. ***Lake Shore Cryotronics Inc****, 575 McCorkle Blvd, Westerville, OH 43082, www.lakeshore.com*

Ultracompact dilution refrigerator

Bluefors has announced its Ultracompact LD dilution refrigerator system, an all-in-one cryogenic measurement system suitable for laboratories with limited space. Dilution refrigerators provide essential ultralow-temperature cooling for quantum devices, such as those used in quantum computing. The system integrates all components necessary for a dilution refrigerator: a Bluefors LD cryostat, its Gas Handling System Generation 2, and a pulse tube compressor. Using multiple dampening solutions to deliver low-vibration operation, the system generates sound levels that are 8–14 dB lower than the standard Bluefors LD. The Ultracompact LD makes installation, setup, and operation simple and efficient with an integrated 19-inch rack that has space for additional measurement electronics and with doors that provide easy access to the gas handling system and pulse tube compressor. ***Bluefors Oy***, *Arinatie 10, 00370 Helsinki, Finland, https://bluefors.com*

Cryogenic epoxy



Master Bond's EP29LPSPND-3, a two-component, nondrip epoxy compound with a paste consistency, can be used for bonding and sealing applications. The system is electrically nonconductive and thermally insulative, with a thermal conductivity of approximately 0.2 W/mK at room temperature. A key performance attribute is its ability to withstand temperature cycling even at cryogenic levels. It is serviceable in the range from 4 K to 394 K. The epoxy features a coefficient of thermal expansion of $45–50 \times 10^{-6}$/°C, a tensile strength of 6000–8000 psi, and a Shore D hardness of 70–80. The system has a volume resistivity exceeding 10^{15} Ω-cm at 75 °F and a dielectric constant of 4.2 at 60 Hz. It has a mix ratio of 100:65 by weight, with a long working life after mixing; a 100 g batch will yield an open time of greater than 5 h at 75 °F. Despite being a paste, it cures clear when applied in thin sections, with a refractive index of 1.56 at 589 nm. ***Master Bond Inc***, *154 Hobart St, Hackensack, NJ 07601, www.masterbond.com*

Turbomolecular pumps for high vacuum

With the Turbovac Mag 2807 iS and 3207 iS Maglev, Leybold introduces two compact turbomolecular pumps in the pumping speed class of 3000 l/s. Equipped with magnetic rotor bearings, the robust, low-vibration pumps are suitable for clean, hydrocarbon-free, high-vacuum conditions. They are suitable for scientific research fields that use accelerators, where clean, vibration-free operation is crucial, and for coating processes, such as the production of glass lenses. They can be used in R&D and industrial fields pertaining to electron beam processes, space-simulation chambers, and beamline applications. According to the company, the pumps offer an excellent ratio of maximum pumping speed to size: The Turbovac Mag iS delivers pumping speeds of up to 3200 l/s at a maximum gas throughput of 26 and 33 mbar l/s for argon and nitrogen, respectively. The lightweight pumps, at less than 60 kg, can be installed in almost any position. Minimal wear allows for maintenance intervals of up to 80 000 h or 10 000 cycles. ***Leybold GmbH***, *Bonner Str 498, 50968 Cologne, Germany, www.leybold.com*

Semiconductor process deposition system

ULVAC's Entron-EXX multichamber deposition system for semiconductor applications builds on its predecessor, the Entron-EX W300, to provide users with an optimized environment for development and mass production. The system integrates various process modules, such as precleaning, heating, cooling, and physical vapor, chemical vapor, and atomic layer deposition. It also features enhanced data collection and analysis capabilities and a scalable design. Real-time processing of large data volumes helps users improve yields, optimize preventative maintenance, and enhance operational efficiency. Because the highly expandable design enables quick module additions and swapping, the system can be easily adapted for evolving needs. The Entron-EXX offers two versatile platform options: single and tandem. The single platform features a single transfer chamber and a simple, space-efficient design. The tandem platform, with two transfer chambers arranged in sequence, is suitable for complex processes that require higher productivity. ***ULVAC Technologies Inc***, *401 Griffin Brook Dr, Methuen, MA 01844, www.ulvac.com* PT

Michael Anand is a meteorologist at the National Weather Service's office in Albuquerque, New Mexico. Since 2022, Anand has provided decision support services for public officials and for organizers of the Albuquerque International Balloon Fiesta.

Ballooning in Albuquerque: What's so special?

Michael Anand

A unique valley and mountain circulation forms a natural route for balloonists to navigate the atmosphere.

The atmospheric conditions in and around Albuquerque, New Mexico, are particularly well suited for hot-air ballooning. Every year in early October, the city hosts the Albuquerque International Balloon Fiesta. Hundreds of hot-air balloonists fly in the sky, as shown in figure 1, and thousands of tourists come to the city to view the spectacle each morning of the nine-day event.

Why Albuquerque? It has a unique wind pattern called the Albuquerque box. Balloon pilots directly control their vertical ascent by heating or venting the air in the balloons, but when it comes to horizontal travel, they're typically at the mercy of the wind. Wind speed and direction can vary with altitude, however, and the pilots can use the variation to their advantage to help control their balloons' flight.

Albuquerque's atmosphere

The clear skies, dry air, and light winds across much of the high desert in the southwest US are especially conducive to a process called radiational cooling. That's when the air near the ground cools after sunset, and heat in the form of long-wave IR radiation—with wavelengths between 8 and 14 μm—escapes into outer space. Without clear skies and the low specific heat of dry air, longwave IR radiation would more easily be trapped by clouds, and the high specific heat of moisture would warm the surface. The light winds in the area also help long-wave IR radiation to escape more easily. Strong winds promote turbulent mixing, which stymies the transport of IR radiation.

As the air cools, it becomes denser. From the mountains around Albuquerque, the cool, dense air flows toward the Rio Grande along the valley floor before eventually developing into a shallow layer across the valley. The flow of cooler air from higher to lower elevations helps form what's called a drainage wind, whose speed is usually no more than about 16 km/h. The drainage wind's path follows the valley's topography: from northern New Mexico's steep and elevated terrain, where surface pressure is high, to the south and through the Rio Grande valley to Albuquerque, where surface pressure is low.

Because of the cooler, denser air that pools in the valley, the air above is warmer and less dense. A vertical temperature profile of the valley of Albuquerque with colder air near the surface and warmer air above would show a temperature increase with height.

The temperature inversion acts like a lid on the air nearest to the ground. Above the temperature inversion, which can be as high as 150 m, the wind usually flows in a different direction from the near-surface drainage wind. The large-scale wind patterns of the atmosphere dictate the direction of the wind above the inversion. To make the Albuquerque box circulation, the wind above the temperature inversion blows

FIGURE 1. HOT-AIR BALLOONS fill the sky of Albuquerque, New Mexico, throughout the year. But in the fall, the especially stable atmospheric conditions promote the Albuquerque box effect, a wind pattern that helps pilots more easily travel around. This 2023 photo is from the Albuquerque International Balloon Fiesta, which attracts thousands of tourists. (Image courtesy of Bennie Bos, Albuquerque International Balloon Fiesta.)

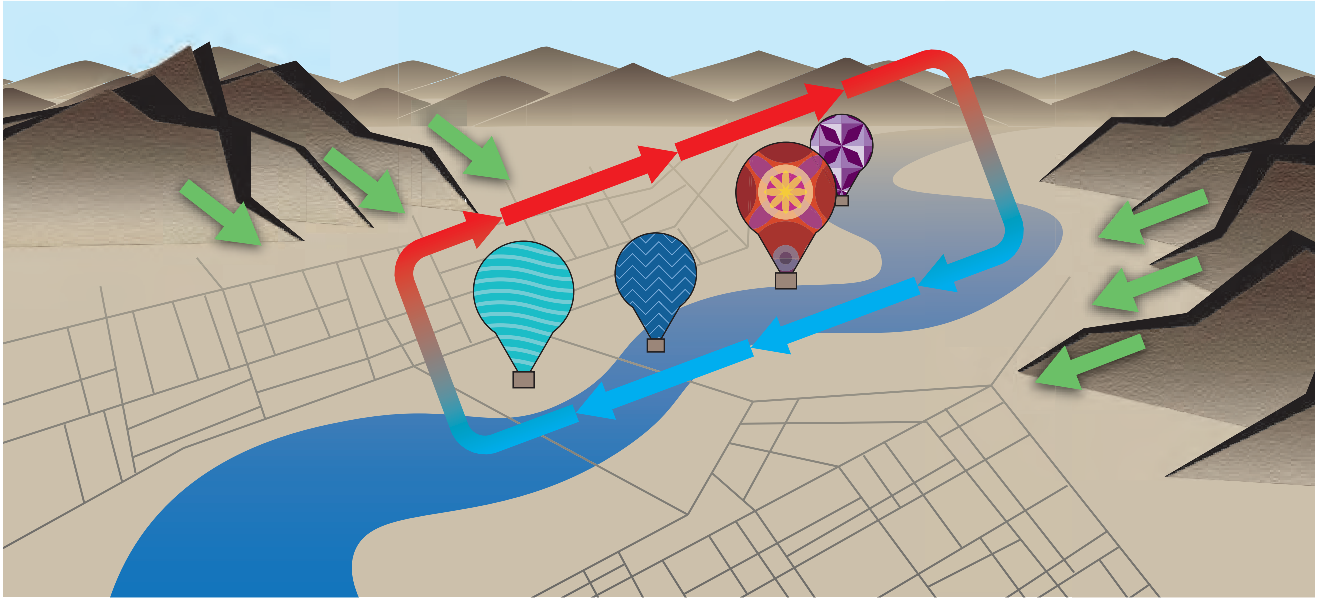


FIGURE 2. THE ALBUQUERQUE BOX is a unique wind circulation in and around the city. When cold, dense air (green arrows) flows from the surrounding mountainous terrain toward the Rio Grande, a northerly drainage wind (blue arrows) develops and moves from higher elevation in the north to lower elevation in the south. In the river valley, temperature increases with altitude, which is called a temperature inversion. Above the inversion, the large-scale wind pattern dictates the box wind's direction. For the Albuquerque box, the large-scale wind (red arrows) travels from south to north. (Image by Freddie Pagani.)

from south to north, which is opposite to the drainage wind in the valley that flows from north to south, as shown in figure 2.

The bounds of ballooning

The Albuquerque box lasts for only a few hours in the morning. When the Sun rises, surface heating produces columns of buoyant air, called thermals, that rise from the surface. The thermals mix with the cooler, denser air, which destroys the inversion. Once the inversion is gone, the higher-altitude wind combines with air closer to the surface and eliminates the Albuquerque box by late morning.

Albuquerque and the rest of the southwest US are located south of the midlatitude region where storm systems and frontal boundaries often develop. But if one of those large-scale atmospheric features forms near Albuquerque, the box circulation cannot occur. A frontal boundary results in a high surface temperature and an anomalous wind gradient. Together, they prevent the development of the valley inversion that's necessary to form the box wind pattern.

The Albuquerque box is most common during the fall season because that is when stable conditions are most common in the atmosphere. In October when the Albuquerque International Balloon Fiesta is held, the pattern occurs about 3 days out of the first 15 days. That may seem infrequent, but the region is still an exceptional place for ballooning. Pilots use the near-surface part of the Albuquerque box to travel south from the takeoff location, which is usually at the fiesta grounds in densely developed Albuquerque. Then, they ascend to higher altitude and use the upper portion of the wind to travel back north, often reaching close to the takeoff location. Such a reliable ballooning experience can't easily be found in other urban areas.

Planning a flight

Pat Chando, a hot-air balloon pilot who works for the National Weather Service office in Albuquerque, shared with me his experience using the wind in the Albuquerque box. Chando says that the winds in the box are not as clearly defined as portrayed in figure 2. The variation is because wind direction is affected by turbulent eddies that form between human structures and buildings. Wind direction is also influenced by different ground surfaces—the wind is more turbulent over concrete and road surfaces, for example, and less turbulent over natural soils and water.

Pilots must also contend with false lift when taking off. The wind at the ground could be completely calm because of frictional effects. Once the balloon rises a few meters, however, it enters stronger wind currents. The northerly drainage wind helps give the balloon false lift—it pushes the top part of the balloon faster than the bottom basket. When the balloon rises farther and goes above that fast current of air, however, it could briefly drop in altitude.

The depth of the inversion layer where the drainage winds are observed is highly variable. It can be as shallow as 3 m or as deep as 150 m. The change in wind direction at the top of the inversion layer is another variable that pilots have to keep in mind and plan for. All the considerations and challenges, however, are what makes ballooning exciting and fun for pilots. Given Albuquerque's box pattern, temperate climate, and generally light and terrain-dominated winds in the valley, ballooning is a year-round activity. The fiesta organizers call the city the Hot-Air Balloon Capital of the World.

Additional resources

- B. Brötz et al., *Bound.-Layer Meteorol.* **152**, 45 (2014).
- J. Schmidli, R. Rotunno, *J. Atmos. Sci.* **69**, 561 (2012).
- S. F. J. De Wekker, M. Kossmann, *Front. Earth Sci.* **3**, 77 (2015).
- T. B. McKee, R. D. O'Neal, *J. Appl. Meteorol. Clim.* **28**, 445 (1989).
- S. Serafin et al., *Atmosphere* **9**, 102 (2018).
- C. D. Whiteman, T. B. McKee, *J. Appl. Meteorol. Clim.* **21**, 290 (1982).

PT

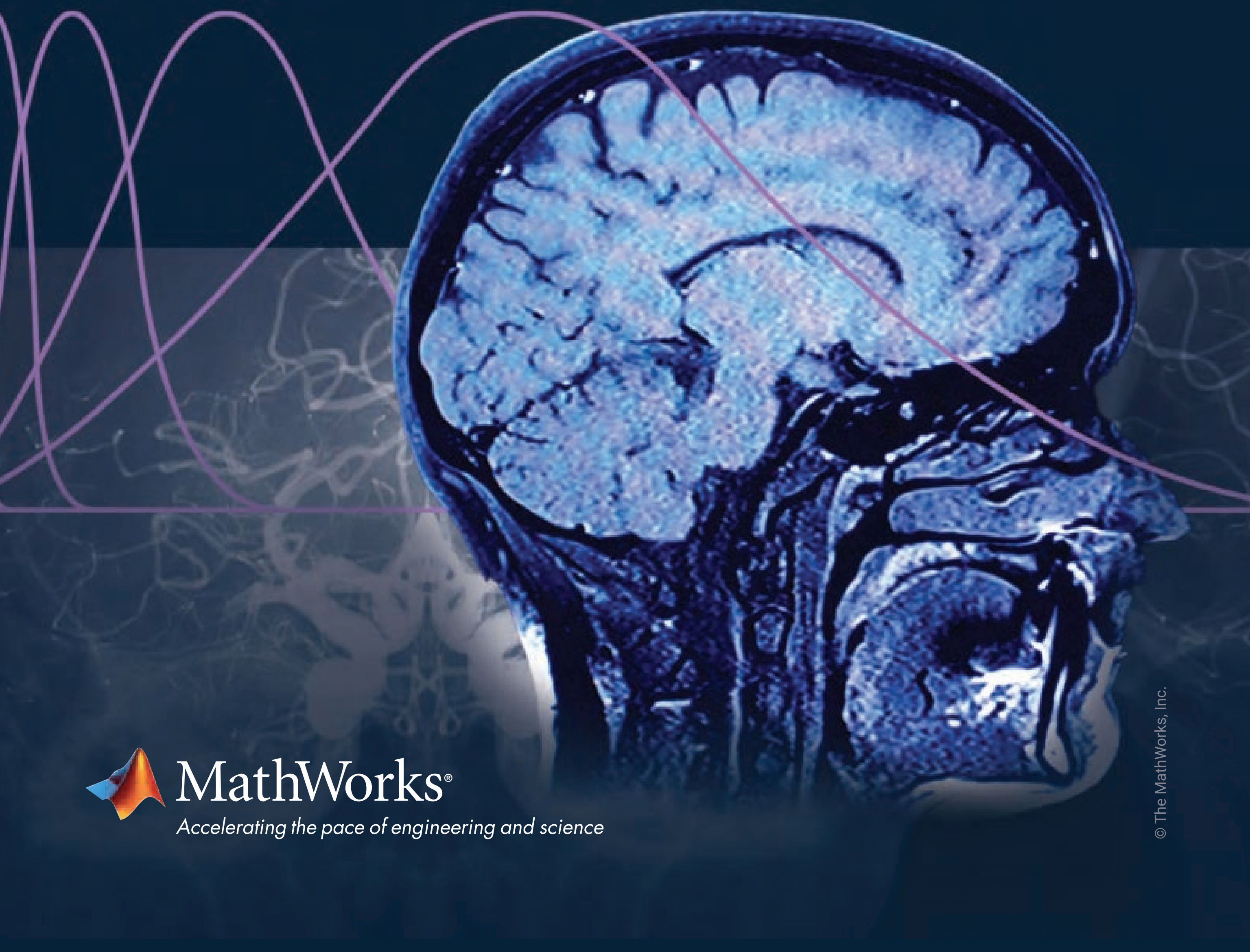
Preventing particle leaks in fusion reactors

Much of the research into harnessing nuclear fusion energy focuses on reactors that use powerful magnets to manipulate ultrahot plasmas. Among the designs is the stellarator. Developed in the 1950s, it consists of a confinement chamber shaped like a twisted doughnut that is wrapped in a set of solenoidal coils, which generate the intense magnetic fields necessary to control plasma. Inside the chamber, deuterium and tritium nuclei fuse to form alpha particles and neutrons.

A major challenge in stellarator design is keeping the alpha particles confined in the chamber, where they maintain the temperature of the plasma and induce further fusion reactions. Escaping alpha particles reduce fusion energy generation and can damage components of the stellarator. But researchers have had trouble locating the holes in the stellarator's magnetic field through which the particles escape. Because it would be computationally infeasible to determine exactly where they are, researchers typically use perturbation theory to make an estimation. A growing amount of evidence indicates that those calculations miss the mark for certain stellarator designs.

Joshua Burby of the University of Texas at Austin and colleagues developed a nonperturbative method to track the loss of alpha particles by training a machine-learning model on data from particle simulations. The diagram shows how the predicted motion of particles in a fusion reactor in the new model (orange and red lines) are closely in line with the motion predicted by time-consuming exact calculations (blue and green lines). Along with helping to refine stellarators, the method could benefit researchers designing tokamaks and other types of magnetic confinement devices. (J. W. Burby et al., *Phys. Rev. Lett.* **134**, 175101, 2025; image courtesy of Max Ruth/University of Texas at Austin.)

—RD



MathWorks®
Accelerating the pace of engineering and science
© The MathWorks, Inc.