PHYSICS TODAY
August 2024 • volume 77, number 8
A publication of the American Institute of Physics

THE QUEST FOR LASER FUSION

Fifty years of dark matter

The wonderland of angstrofluidics

The smallest drops of acid


the SPS Observer
Volume LVII, Issue 2
FALL 2023
Presented by GradSchoolShopper,
the Society of Physics Students,
and the American Institute of Physics
THE GRAD SCHOOL ISSUE
READ NOW!
gradschoolshopper.com/magazine
Grad Programs Look For
+ How to Build a Strong CV and Résumé
+ How to Get Great Letters of Recommendation
Now You Know: Grad School Edition
SPS
GradSchoolShopper
presented by
AIP | Empowering Physical Scientists

Find out more

Unlock the Full Spectrum.
From DC to 8.5 GHz.

Your Application. Measured.

Nanotechnology & Materials Science

Optics & Photonics

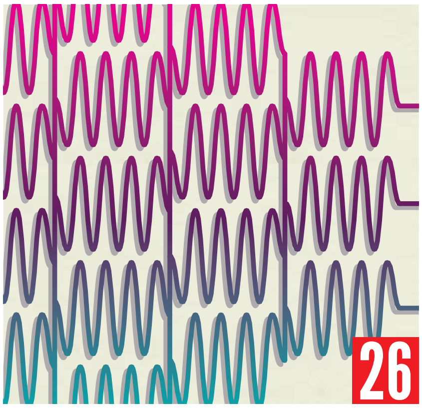
Impedance Analysis

Scanning Probe Microscopy

Sensors

Failure Analysis & Semiconductors

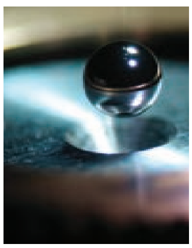
Zurich Instruments



PHYSICS TODAY

August 2024 | volume 77 number 8

FEATURES



ON THE COVER: When a laser heats and compresses this 2 mm capsule to a high-enough temperature and density, the deuterium–tritium fuel that's inside will ignite. In recent experiments, the resulting fusion reaction has produced more energy than was put into it. With the fundamental physics better understood, researchers are now working toward developing fusion for commercial use. To learn more, see the article by Stefano Atzeni and Debra Callahan on **page 44**. (Courtesy of Lawrence Livermore National Laboratory/CC BY-NC-SA 4.0.)

Recently on PHYSICS TODAY ONLINE

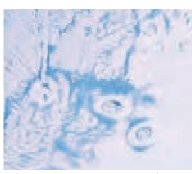
www.physicstoday.org

© RAISSA PAGE

A historic protest

Before disbanding in 2000, a group of antinuclear protesters—all of them women—maintained an encampment near a UK air force base for nearly two decades. PHYSICS TODAY's Ryan Dahn examines the Greenham Common Women's Peace Camp through the lens of the participants' self-published newsletters.

physicstoday.org/Aug2024a

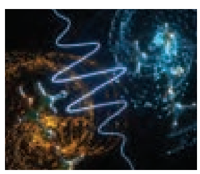


COPERNICUS SENTINEL DATA/R. DELL

Ice-shelf slush

In satellite images of Antarctica's ice shelves, pools of meltwater appear a vivid blue—but areas of waterlogged slush can be confused with other features. With a new machine-learning algorithm, researchers can better identify the slushy spots and estimate the danger the meltwater poses to the ice shelves.

physicstoday.org/Aug2024b



ELLA MARU STUDIO INC

Chiral dynamics

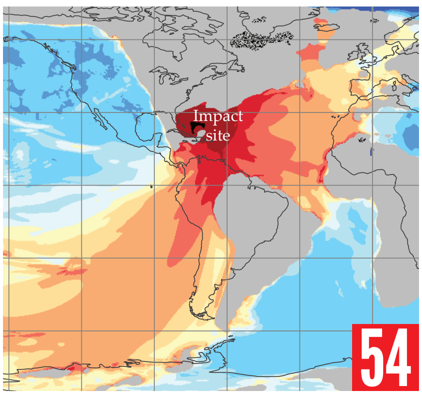
A molecule excited with a light pulse that lasts just a few femtoseconds can show a response that's consistent with the molecule's mirror-image counterpart, a new study reveals. The research offers insight into how the motion of electrons in chiral molecules affects reactivity.

physicstoday.org/Aug2024c

PHYSICS TODAY (ISSN 0031-9228, coden PHTOAD) volume 77, number 8. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Periodicals postage paid at Huntington Station, NY, and at additional mailing offices. POSTMASTER: Send address changes to PHYSICS TODAY, American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Views expressed in PHYSICS TODAY and on its website are those of the authors and not necessarily those of AIP or any of its member societies.

PHYSICS TODAY

www.physicstoday.org



DEPARTMENTS

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

PhysicsToday
Physics Today
@physicstoday

The American Institute of Physics is a federation of scientific societies in the physical sciences, representing scientists, engineers, educators, and students. AIP offers authoritative information, services, and expertise in physics education and student programs, science communication, government relations, career services, statistical research in physics employment and education, industrial outreach, and history of the physical sciences. AIP publishes Physics Today and is also home to the Society of Physics Students and to the Niels Bohr Library and Archives. AIP owns AIP Publishing, a scholarly publisher in the physical and related sciences.

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

Five-molecule water clusters have liquid-like properties

With theory-guided rotational spectroscopy measurements—and some help from nuclear spin—researchers can tell whether a hydrogen chloride molecule in a tiny amount of water is dissolved.

WATER DROPLETS come in various sizes. But how small can they be before they stop acting like liquids? (Image by Aathavan jaffna/Wikimedia Commons/CC BY-SA 3.0.)

How big is a drop of water? That's not an unanswerable question, like "How long is a piece of string?" A single H_2O molecule isn't a drop. As more and more molecules cluster together, there must be some threshold at which they start displaying behaviors associated with the condensed phase.

Nor is the question a mere philosophical curiosity. How many water molecules make a drop is relevant to the chemistry of such environments as Earth's atmosphere and interstellar space, where molecules are typically found in isolation or in small clusters. It's also relevant to computational chemistry: Simulating molecules takes a lot of computing power, so modelers want to know how many molecules they really need to reproduce liquid water's properties.

How does one even tell whether water is acting like a bulk liquid? One thing that condensed-phase water can do, but single molecules can't, is dissolve other substances; as a paradigmatic case, the research community has settled on hydrogen chloride. High school chemistry students learn that HCl is a so-called strong acid: When dissolved in water, its molecules always break up into H^+ and Cl^- ions, with the former latching onto water molecules to make H_3O^+. On its own, though, HCl is a covalently bonded molecule. So how much water is needed to split it into an ion pair?

At the German Electron Synchrotron (DESY), Melanie Schnell and colleagues are now shedding some light on that question.[1] She and her postdocs Fan Xie, who worked on the experimental side of the study, and Denis Tikhonov, who handled the theory, have conducted one of the most detailed and thorough studies yet on HCl microhydration. They found that in clusters of four water molecules, HCl is intact, but with five water molecules, it splits into a pair of ions. Although "five" isn't a universal answer to "How many water molecules make a drop?"—different amounts of water might be needed under different conditions and to dissolve different substances—the work is a step toward demarcating the boundary between the molecular and condensed-phase worlds.

Standing on shoulders

Water-cluster structures can be hard to elucidate. There's no way to zoom in and snap a photo of a single cluster to see whether an HCl molecule is dissociated. Some techniques come close to directly visualizing the atom-by-atom structures of molecules (see Physics Today, December 2017, page 22, and May 2022, page 12), but they typically require many identical specimens, so they're not suitable for studying water clusters. Clusters can be made en masse by spraying HCl and H_2O together in a jet of a carrier gas, such as neon. But there's no control over the number of molecules in each cluster, let alone their arrangement, so the resulting clusters come in an enormous number of different structures.

Importantly, though, "enormous" is not "infinite." In the cold environment of the neon jet, clusters tend to settle into structures that are local energy minima. Just like molecules, those cluster structures can be calculated with density functional theory (DFT) or similar methods. And just like molecules, the clusters have discrete rotational and vibrational quantum states, which means they can be probed spectroscopically.

The strongest spectral lines are those of single molecules and dimers: The molecules can arrange themselves in fewer ways, so each structure creates a larger share of the signal. As researchers figured out the structures of larger clusters by matching the spectral lines with their computed DFT predictions, those lines could be subtracted from the spectrum. What was left was a forest of weaker and more numerous lines from larger and more complicated clusters.[2]

It's been a decades-long quest by many research groups to untangle the IR (vibrational) and microwave (rotational) spectra of water clusters, with and without HCl or other guest molecules, and Schnell acknowledges that her group's work builds on previous efforts. "In particular, Zbigniew Kisiel's studies on smaller HCl–water clusters significantly eased the challenges in our analysis of larger clusters," she says. "Our work would have been much more difficult and tedious without those pioneering efforts."

Spin splitting

Several factors allowed Schnell and colleagues to progress from studying clusters with three or four water molecules

FIGURE 1. IN A CLUSTER of water molecules, a hydrogen chloride molecule can dissociate into a separated ion pair (left) or a contact ion pair (center), or it can remain undissociated (right). In four-water clusters, such as the ones shown here, only the intact HCl structures were observed. But with five or more water molecules, the H–Cl bond can break. (Adapted from ref. 1.)

to those with five or more. First, they used a high-resolution microwave spectrometer, which sampled the spectral range from 2 to 8 GHz with a resolution of 25 kHz, so they could distinguish spectral lines that previous lower-resolution IR studies could not.[3]

Second, their DFT search for cluster structures included not just the energy-minimizing structures themselves but also the pathways between those minima. Many of the structures that the researchers calculated were similar in energy and had similar predicted rotational spectra. Only by understanding how the structures might interconvert could the researchers figure out which structures were actually showing up in the spectrum.

Third, they exploited an interaction between rotational quantum states and nuclear spin. Both of chlorine's stable isotopes, ^{35}Cl and ^{37}Cl, have electric quadrupole moments, which impart a hyperfine splitting to spectral lines. But the splitting is dampened when the nucleus is surrounded by a complete electron shell, as in a Cl^- ion. When Cl is part of an H–Cl covalent bond, on the other hand, the spin is only partially blocked by electrons, and the splitting is much larger. By measuring the magnitude of hyperfine splitting, Schnell and colleagues found valuable clues to figure out which cluster structure they were looking at—and, crucially, whether the HCl molecule was intact or dissolved.

None of those factors by itself is new. The nuclear-spin effect, in particular, was described by Charles Townes and Benjamin Dailey back in 1949.[4] But they'd never been applied together to the H_2O–HCl cluster system. "Before our studies, the search for HCl dissociation with a few water molecules was almost there," says Schnell. "We just gave it a final kick."

Four, five, six, seven

Acid–water clusters come in three types, as shown in figure 1. In some, the HCl molecule is fully dissociated into a separated ion pair: The Cl^- ion, shown in green, sits on one side of the cluster, and its H^+ partner (white) is attached to an H_2O molecule on the other side. Some are what the researchers call contact ion pairs: The H^+ ion is joined to a water molecule but still adjacent to the Cl^- ion. And in others, the HCl molecule is undissociated.

Schnell and colleagues didn't observe any separated ion pairs. At first, they were surprised and disappointed, but when they looked more closely at their DFT calculations, the result made sense. All of the

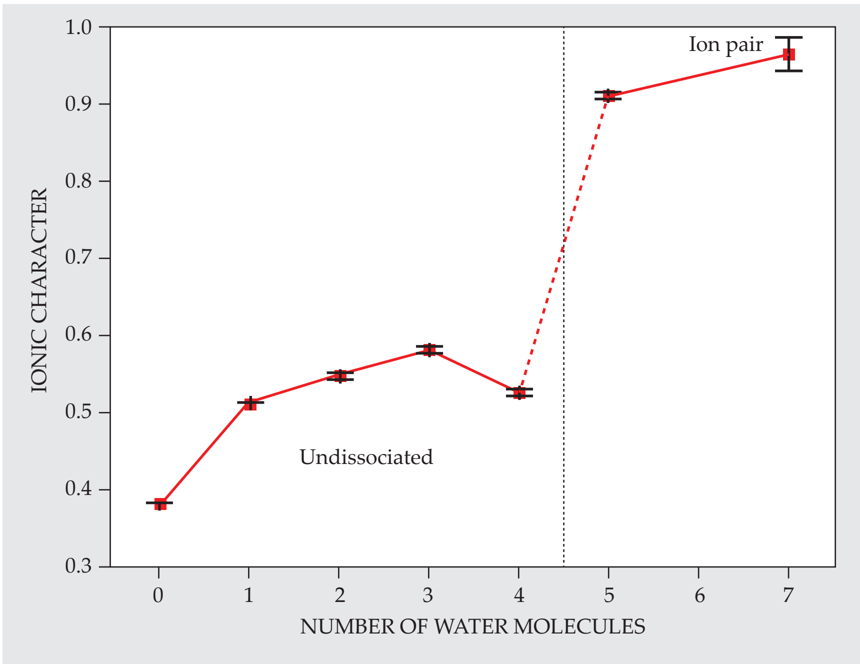


FIGURE 2. BETWEEN FOUR AND FIVE WATER MOLECULES in a cluster, there's an abrupt transition in which a hydrogen chloride molecule dissociates from a covalently bonded molecule into a pair of ions. The chlorine's ionic character is quantified on a scale from 0 to 1 by the Cl nuclear spin's effect on hyperfine splitting. When the Cl atom is part of a molecule, the splitting is relatively large and the ionic character is low. But in a free Cl^- ion, the nucleus is shielded by a complete electron shell, which dampens the splitting almost to zero. (Adapted from ref. 1.)

energy barriers to forming those structures were too high for clusters to surmount in the cold environment of the neon jet.

The analysis therefore came down to distinguishing contact ion pairs from undissociated structures, and the Townes–Dailey nuclear-spin effect was crucial. The two classes of structures can be hard to differentiate from the arrangements of atoms alone, because shifting an H atom a fraction of an angstrom from an HCl molecule to an H_3O^+ ion might not have much of an effect on the cluster's rotational energy levels. But it makes all the difference to the hyperfine splitting.

From the splitting that they measured, the DESY researchers calculated the clusters' ionic character on a scale from 0 (for a nonpolar covalent bond, as in a Cl_2 molecule) to 1 (for a Cl^- ion). As figure 2 shows, a zero-water HCl cluster—that is, just an isolated HCl molecule—already has an ionic character of almost 0.4. That's because the H–Cl bond is polar, with more of its electron density on the Cl than on the H. For clusters with one to four water molecules, the ionic character doesn't change much. But for five- and seven-water clusters, there's a clear jump: The Cl is almost completely ionic.

What about clusters with six water molecules? The researchers have every reason to think that those will also form contact ion pairs, but they're still working on the spectral assignments. Perhaps counterintuitively, six-water clusters are harder to study than seven-water clusters, because they tend to have lower symmetry. With five or seven water molecules, the Cl^- and H_3O^+ ions can sit in the middle of the cluster, with the rest of the H_2O molecules arranged in pairs or trios symmetrically around the edges. With six water molecules, there's no such option, and the lower-symmetry clusters give rise to more complicated spectra.

The more pressing questions are the ones the researchers don't know the answers to. Can the clusters be made to form separated ion pairs, not just contact ion pairs, either by adding more water molecules or changing the experimental conditions? Do other acids behave the same way as HCl? Spectroscopic study of acid–water clusters is not going to get any easier, as the simplest remaining structures have their spectral lines assigned. But with the right combination of experimental and theoretical tricks—including the Townes–Dailey nuclear-spin effect—further progress may be possible. "Nuclear spin is really cool," says Xie. "It can reveal molecular properties that are difficult to assess in other ways."

Johanna Miller

References

1. F. Xie, D. S. Tikhonov, M. Schnell, *Science* **384**, 1435 (2024).
2. See, for example, Z. Kisiel et al., *J. Phys. Chem. A* **104**, 6970 (2000).
3. D. Schmitz et al., *J. Mol. Spectrosc.* **280**, 77 (2012).
4. C. H. Townes, B. P. Dailey, *J. Chem. Phys.* **17**, 782 (1949).

How to make a midsize black hole

Detailed simulations of star formation show that runaway collisions in a giant molecular cloud could produce very massive stars that are precursors to intermediate-mass black holes.

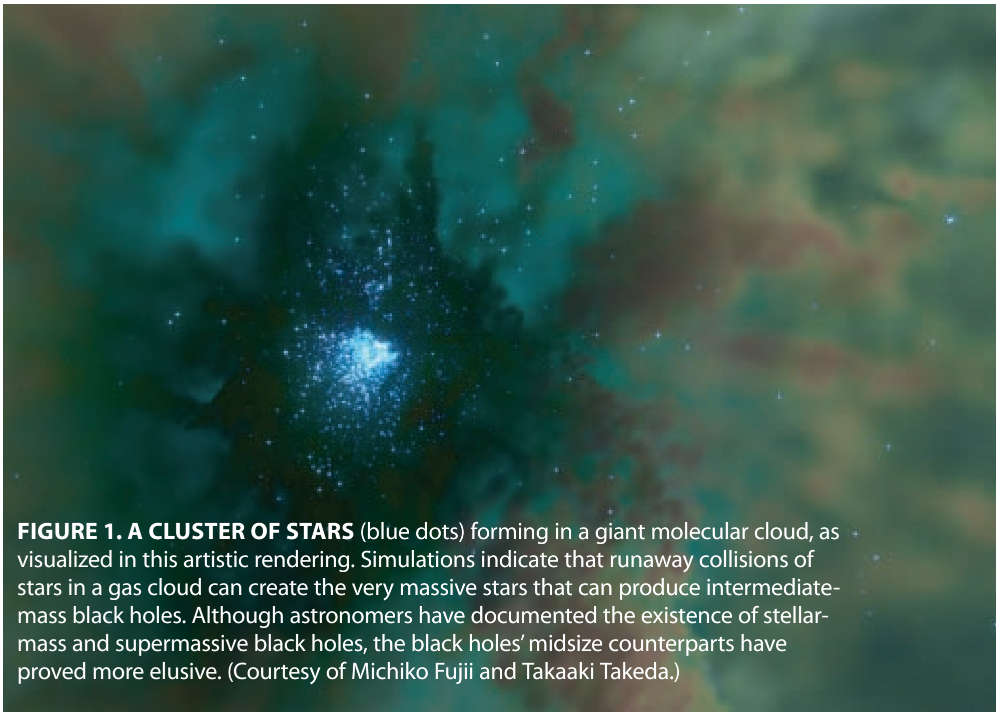


FIGURE 1. A CLUSTER OF STARS (blue dots) forming in a giant molecular cloud, as visualized in this artistic rendering. Simulations indicate that runaway collisions of stars in a gas cloud can create the very massive stars that can produce intermediate-mass black holes. Although astronomers have documented the existence of stellar-mass and supermassive black holes, the black holes' midsize counterparts have proved more elusive. (Courtesy of Michiko Fujii and Takaaki Takeda.)

Found at the center of galaxies, supermassive black holes have the mass of more than 100 000 suns. The largest ones tip the scales at billions of solar masses. Stellar-mass black holes, widely observed in binary systems with a partner star, are typically less than 50 solar masses. But observations of black holes with masses somewhere in between have been relatively sparse. That gap has long puzzled astronomers, who have wondered how supermassive black holes could get so enormous without any intermediate-mass black holes to grow from.

Although astronomers generally agree that stellar-mass black holes form from the collapse of individual stars, the formation mechanisms for intermediate-mass and supermassive black holes are not as well understood. Ideas include the formation and collapse of very massive stars, the merging of smaller black holes, and the accumulation of mass once a black hole has formed. Findings show that some of the largest observed supermassive black holes formed when the universe was less than a billion years old,[1] but it's unclear how they grew so quickly.

New star-by-star simulations by the University of Tokyo's Michiko Fujii and colleagues lend support to one possibility: Stars that grow in dense molecular clouds can merge with the speed and efficiency needed to grow into very massive stars that will eventually collapse to form intermediate-mass black holes.[2] The results suggest that astronomers looking for the elusive midsize black holes should continue to target dense groups of stars known as globular clusters, which are found in the galactic halo of the Milky Way.

Merging codes

Fujii and her team had set out not to simulate the formation of intermediate-mass black holes but rather to investigate the growth of globular clusters from a molecular gas cloud. To do that, the researchers brought together two disparate codes. One simulates the process of star formation from molecular gas clouds, and the other, known as an *N*-body code, models the gravitational interactions of individual stars.

By making improvements to the algorithms in the *N*-body code, Fujii's team could run simulations that contain millions of individually resolved stars. Previous *N*-body codes couldn't include so many individual stars because of computational limitations, so researchers have relied on simplifying assumptions to model star clusters: grouping many stars together as individual particles or using Monte Carlo simulations that contain other simplified parameters.

Fujii and colleagues integrated the more sophisticated *N*-body code with a model of star formation in gas clouds, as illustrated in figure 1. The new code enabled the researchers to simulate the birth of the Orion Nebula and, with the latest study, dense star clusters from their inception. "That they include both gas and stars and have a realistic number of stars is really impressive," says Nathan Leigh of the University of Concepción in Chile. "This brings simulations into a more realistic regime."

Accumulating mass fast

Fujii and colleagues' simulations begin with a gas cloud that contains 10^5–10^6 solar masses worth of matter. The high-density gases form filaments and clumps that eventually collapse to form stars that proceed to grow, interact, and merge amid the cloud.

Generally, as stars grow larger, they generate stronger stellar winds that cause them to lose mass. The competition between mass loss and mass accumulation can thus limit their growth. But the simulations show that the gravitational potential from the dense gas cloud around the stars keeps the star cluster compact. So as more stars merge, mass accumulates faster than it is lost. Eventually, stars can form that are more than 1000 solar masses, much more massive than any stars ever observed. As shown in figure 2, those very massive stars, as they're known, gain most of their mass in under 100 000 years. "We expected some collisions of stars to proceed inside globular clusters, but it was surprising to us that we formed such a big star," says Fujii.

Though such massive stars have never been observed, their existence has been predicted by theorists. There are many reasons such enormous stars may not have been sighted—they could have a cool surface that makes them hard to spot in the sky, or they could be short-lived before they collapse into black holes. Other theoretical work has estimated that such a star would collapse into an intermediate-mass black hole after about 2 million years. Fujii and colleagues used stellar-evolution calculations to estimate mass loss from a very massive star over that duration, as shown in figure 2. But because very massive stars remain only theoretical, the calculations of how they would evolve are a best guess.

"Nobody knows how a 1000 solar mass star evolves, what is the nuclear burning process in these extraordinary massive objects," says Simon Portegies Zwart of Leiden University.

To assess whether the growth of very massive stars would be limited to the early universe or could be ongoing, Fujii and colleagues explored different values of metallicity, the proportion of elements heavier than helium. Stars that formed earlier in the universe's history have lower metallicity. Generally, stars with higher metallicity generate stronger stellar winds that reduce their mass, so lower metallicity is considered favorable for the formation of very massive stars.

The simulations contained varied metallicities, from levels similar to those observed in Milky Way globular clusters to a value five times as high. Models with lower metallicity produced the largest stars in the simulations, but runaway collisions still produced very massive stars in the high-metallicity scenarios. That finding suggests that very massive stars could still be forming today.

Seeing in the dark

What Fujii and colleagues found in their simulations supports results from simpler models published decades ago that identified globular clusters as potential sites for the growth of intermediate-mass black holes.[3] The new result encourages ongoing exploration of those clusters, but obstacles to finding midsize black holes remain. Observational challenges that are unique to the intermediate-mass range may partially explain their seeming absence. "It's strange to have a gap of four orders of magnitude," says Jillian Bellovary of Queensborough Community College in New York. "It doesn't mean they aren't there. We haven't discovered the stuff in the middle because it's hard."

Being black objects against a black background, black holes can be difficult to observe. Orbital dynamics are one reliable way to spot them. Supermassive black holes at the center of galaxies are more easily identified because of their enormous size and also because lots of stars orbit them. Stellar-mass black holes frequently form in binary systems, where their partner stars' mutual orbits can be observed. In both of those systems, frictional heating from the siphoning of matter into the black hole also generates observable x rays and radio waves.

Intermediate-mass black holes are less likely to have easily discerned orbiting stars that would make their presence known. The high density of stars in globular clusters makes it hard to decipher individual orbits, although a recent study may have broken through that barrier: An analysis of two decades of data from the *Hubble Space Telescope* reveals seven stars at the center of Omega Centauri, a Milky Way globular cluster, that are moving so fast that they should escape the cluster unless they are orbiting a black hole of at least 8200 solar masses.[4]

One of the clearest to-date observations of an intermediate-mass black hole took place in 2019, when the merger of two stellar-mass black holes produced a black hole of about 142 solar masses. Nine solar masses' worth of energy radiated from that merger as gravitational waves, which were detected by the twin detectors of the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the US and the Virgo observatory in Italy.

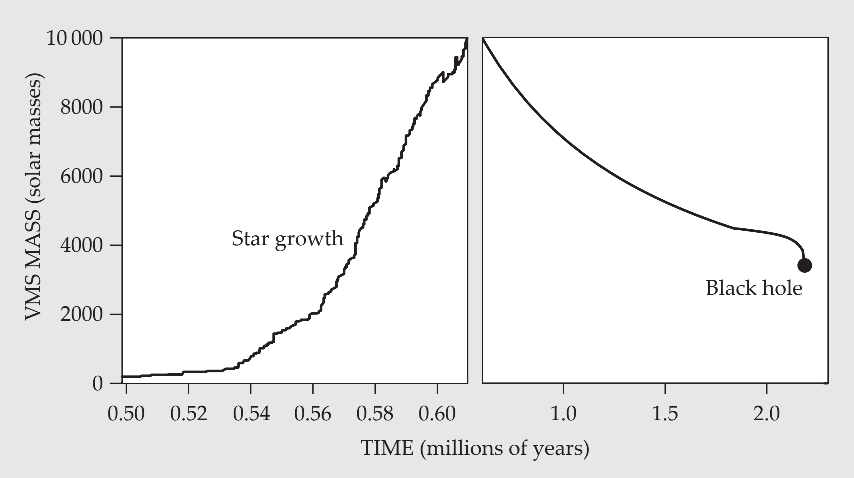


FIGURE 2. VERY MASSIVE STARS (VMSs) can be produced by runaway collisions of smaller stars in a globular cluster. This detailed simulation indicates that a very massive star accumulates most of its mass in less than 100 000 years. After formation stops, such a large star could evolve and collapse into an intermediate-mass black hole over a few million years. (Adapted from ref. 2.)

The LIGO–Virgo observation shows that the collisions of stellar-mass black holes represent one viable method for producing black holes at the smaller end of the intermediate-mass range. Such collisions of stellar-mass black holes, not to mention collisions of stars, are most likely to occur in dense star clusters. But a merger of black holes of different sizes produces asymmetrical gravitational waves that can kick the resulting black hole out of the star cluster, where it is less likely to merge with other objects.

Tidal dissipation allows stars to absorb more of the energy from mergers, so stellar collisions don't produce the same kind of kick that merging black holes do. Thus a very massive star that collapses into an intermediate-mass black hole has been favored as a formation mechanism over the merging of smaller black holes. But there's no reason both can't happen. Once a black hole is massive enough, a merger with a small black hole no longer produces a kick large enough to displace the merged black hole.

The size scale of LIGO and Virgo limits the frequencies of gravitational waves that they can pick up, so they can't detect mergers of black holes at the larger end of the intermediate-mass range. The Laser Interferometer Space Antenna (LISA) mission—led by the European Space Agency in collaboration with NASA and the LISA Consortium—aims to fill that gap (see PHYSICS TODAY, July 2010, page 14). Scheduled to launch in the mid 2030s, LISA is made of three spacecrafts that will be separated by millions of kilometers and will trail Earth in its orbit like a cartwheeling triangle. By sending lasers between its three crafts, LISA will use the same interferometry technique that LIGO and Virgo do, but it will be able to detect gravitational waves at the low frequencies needed to find larger collisions.

As the hunt for more intermediate-mass black holes continues, Fujii and colleagues' simulations confirm that globular clusters are good targets. Fujii plans to continue expanding the size of the star-by-star simulations. Her next goal is to simulate the first star clusters that formed in the universe with almost no metallicity because they may have the potential to form even more massive stars.

Laura Fattaruso

References

1. E. Bañados et al., *Nature* **553**, 473 (2018).
2. M. S. Fujii et al., *Science* **384**, 1488 (2024).
3. S. F. Portegeis Zwart, S. L. W. McMillan, *Astrophys. J.* **576**, 899 (2002).
4. M. Häberle et al., *Nature* **631**, 285 (2024).

Levitating beads reveal radioactivity

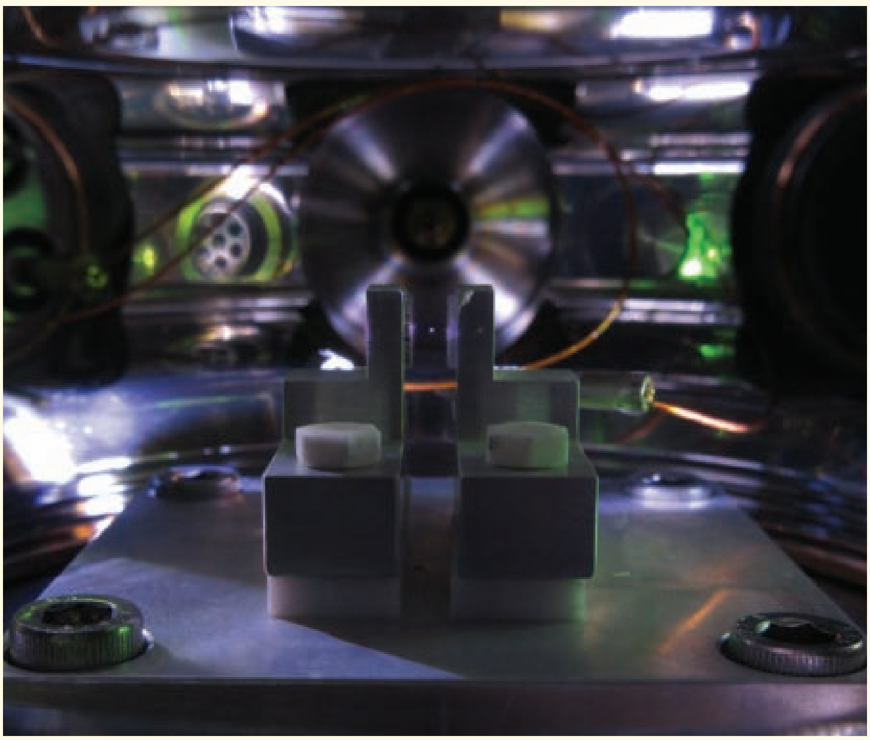
A new technique for studying nuclear decays relies not on complicated particle detectors but on Newton's third law.

The beads under study by David Moore's group at Yale University barely move as they hover inside a vacuum chamber. A focused laser confines the 3-micron-wide silica spheres and cools them to minimize thermal jitters. Yet every now and then, the beads jolt 10 or so nanometers. The cause of each abrupt kick: the recoil of a nucleus in the bead that underwent radioactive decay. The new demonstration offers a blueprint for measuring radioactivity without the need to detect often-elusive decay products.

The technique works by measuring the momentum of a radioactive nucleus that recoils in one direction upon emitting a particle in the other. The researchers embedded solid beads with radioactive lead nuclei and loaded individual beads into the optical trap. Observing each bead continuously for up to three days, Moore and colleagues recorded a series of sudden jumps. The events could be explained by nuclei that emitted alpha particles, recoiled, and then transferred their momenta after crashing into neighboring atoms in the bead. By also tracking the overall charge of each bead, the researchers were able to confirm that changes in charge accompanied those in position.

The technique should work for all types of radioactive decays: Nuclei recoil regardless of whether their emissions include alpha particles, which are generally easy to measure via traditional detectors, or neutrinos, which are not. That thoroughness could prove valuable for researchers probing fundamental physics or investigating the precise compositions of nuclear materials. But first, they will have to improve the technique's sensitivity to enable the detection of lower-energy radioactive processes, such as beta emission and some hypothesized rare decays. Moore's group plans to spot the recoils associated with those processes by using smaller, less-massive spheres. (J. Wang et al., *Phys. Rev. Lett.* **133**, 023602, 2024.)

Andrew Grant



A SMALL SILICA SPHERE hovers nearly motionless in a vacuum chamber. (Courtesy of Yale Wright Lab.)

Long-necked single-celled predator has mastered an unusual origami fold

The tiny organism is nature's version of a bendy straw.

Imagine if you could stretch your arms or neck by a third of the length of a football field in just seconds to grab a snack. The single-celled organism *Lacrymaria olor* can do the scaled-down version of just that. Normally about 60 μm long, the cell regularly extends a necklike appendage up to 1.2 mm to hunt prey.

Single-celled shape-shifters aren't uncommon, but *L. olor* stands out for its speed and reversibility: The full extension–retraction cycle takes 30 seconds and can be repeated more than 20 000 times in the cell's lifetime. The cell isn't encumbered by bones or a spinal cord, but it still has a cytoskeleton and an unstretchable cell membrane, which would seem to preclude such rapid shape change. The neck extends and contracts far faster than the cell can construct new membrane material. Where does all the extra membrane come from? And what physical processes govern the change?

Stanford University's Eliott Flaum and Manu Prakash now shed light on those questions. Their fluorescence-microscopy images in figure 1 show that *L. olor*'s cytoskeletal filaments are wrapped helically around the cell. When they looked more closely, they saw something extraordinary: In the gaps between the filaments, the cell membrane was doubled over on itself in pleats, which unfolded as the neck extended and refolded when it contracted.

The pleated membrane's mechanical properties can be understood in terms of curved-crease origami. Unlike more typical straight folds, which can be partially unfolded to any angle with no energetic penalty, curved creases are often bistable, and they pop between completely folded and completely unfolded states. The phenomenon can be seen in the circular pleats in a bendy drinking straw, shown schematically in figure 2a. But whereas bendy-straw pleats pop one by one at random, *L. olor*'s helical pleats pop all at once, but only partway down the length of the cell, as shown in figure 2b.

Curved-crease-origami energetics go a long way toward explaining how *L. olor* can extend and retract its neck so quickly: Cilia at one end of the organism get the unfolding started, and mechanics and geometry take care of the rest. *L. olor* is not the first unusual single-celled organism that Prakash and his group have studied (see PHYSICS TODAY, September 2019, page 22), and it likely won't be the last. As he notes, the microbiological world has many thousands more species whose unique behaviors remain to be studied. (E. Flaum, M. Prakash, *Science* **384**, eadk5511, 2024.)

Johanna Miller

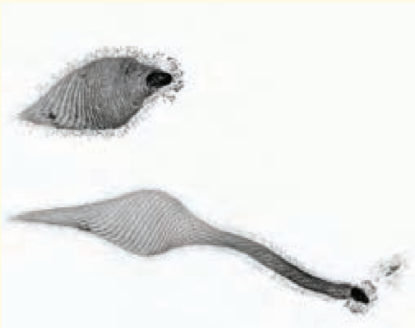


FIGURE 1. THE PROTIST *LACRYMARIA OLOR* can extend its neck to some 20 times its body length and reversibly retract it in seconds. Essential to its extraordinary ability are the helical cytoskeletal filaments, shown in these color-reversed fluorescence-microscopy images, and the cell-membrane pleats between them. (Courtesy of Eliott Flaum and Manu Prakash.)

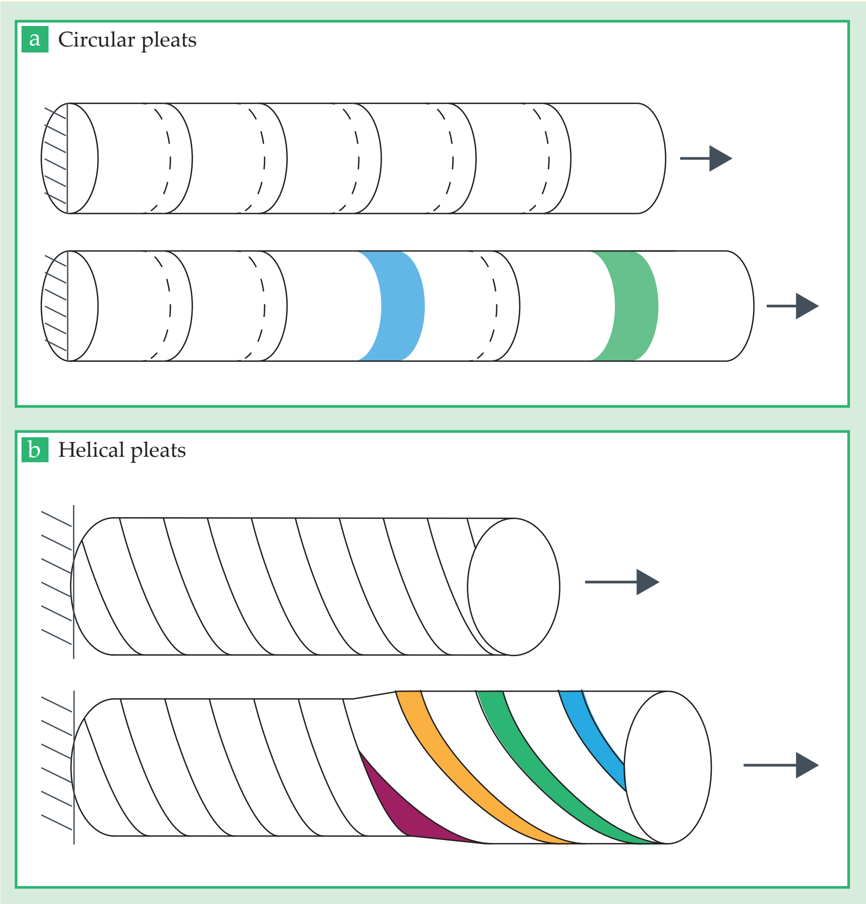


FIGURE 2. BENDY STRAWS AND PREDATORY CELLS both rely on the mechanics of curved creases. The circular pleats of a flexible straw, shown in **(a)**, pop from folded to unfolded one by one. The helical pleats of *Lacrymaria olor*'s cell membrane, shown in **(b)**, behave slightly differently. In their paper's supplementary information, Eliott Flaum and Manu Prakash include detailed instructions for readers to build their own paper model of the spiral-pleated cell. (Adapted from E. Flaum, M. Prakash, *Science* **384**, eadk5511, 2024.)

Rapidly strained metals strengthen when heated

In high-temperature microballistic experiments, rapidly strained copper is about as strong as steel.

Metals, when heated, soften and become easier to deform, and once cooled they harden again. As temperature rises, atomic defects move more readily through the crystal lattice and can break its atomic bonds. Things can get weird under extreme conditions—namely, under high rates of strain. When an object is rapidly deformed over a short time period, the atoms don't have a lot of time to rearrange around the defects. At the macroscopic scale, that means the metal doesn't soften. In a high-strain-rate regime, energetic crystal-lattice vibrations become the dominant obstacle for defects and oppose their motion around the crystal lattice. The less the defects move, the stronger the metal is.

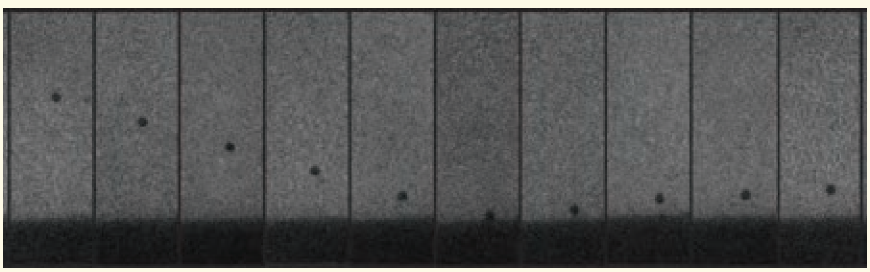
Until recently, researchers had no way to test the theoretical research on how high strain rates affect a metal's strength. The typical approaches can measure strain rates of up to 10^4/s. That means that the object would lengthen to 10 000 times its original length in one second. But in practice, the stress that's applied to a material in high-strain-rate experiments is sustained for just a few nanoseconds.

Although a strain rate of 10^4/s is high—on the order of what's experienced by an object impacted by a meteorite—it's still a few orders of magnitude below the range in which metals should start to behave counterintuitively. Now, because of advances in laser-driven microballistic systems, Christopher Schuh of Northwestern University and graduate student Ian Dowding of MIT have experimentally confirmed that at the high strain rates of 10^6/s–10^9/s, copper and other metals strengthen as temperature increases.

In laser-driven microballistic experiments, fast-moving, micron-sized particles strike a metal target to generate high strain rates in the metal. With MIT postdoc Alain Reiser, Schuh designed a new testing setup that could withstand the high temperatures at which the unusual metal behavior can be observed. Starting at 0 °C and increasing the temperature of the metal by several hundred degrees, Dowding and Schuh measured a copper target's deformation and its strength—the applied stress necessary to deform it—as it was bombarded with 12 μm particles of alumina.

The snapshots shown below, taken with a high-speed camera, illustrate a characteristic particle trajectory. In their experiment, Dowding and Schuh found the sought-after evidence: In a regime of high strain rates, the strength of copper increased by 30% for a 157 °C rise in temperature. The researchers found the same results in gold and titanium as they did in copper, and all pure metals should exhibit the effect. Many applications expose metals to high temperatures and high strain rates, so the findings may be relevant for various industrial processes, including high-speed metal machining, metal additive manufacturing, and sandblasting. (I. Dowding, C. A. Schuh, *Nature* **630**, 91, 2024.)

Alex Lopatka



A SIDE VIEW of the flight of a micron-sized particle impacting and rebounding from a copper target. The time between each image is 500 ns, and the particle's inbound velocity is 80 m/s. Using the particle's rebound velocity and the impact crater's volume, researchers determined that under certain conditions, rapidly strained copper strengthens as temperature increases. (Courtesy of Ian Dowding.)

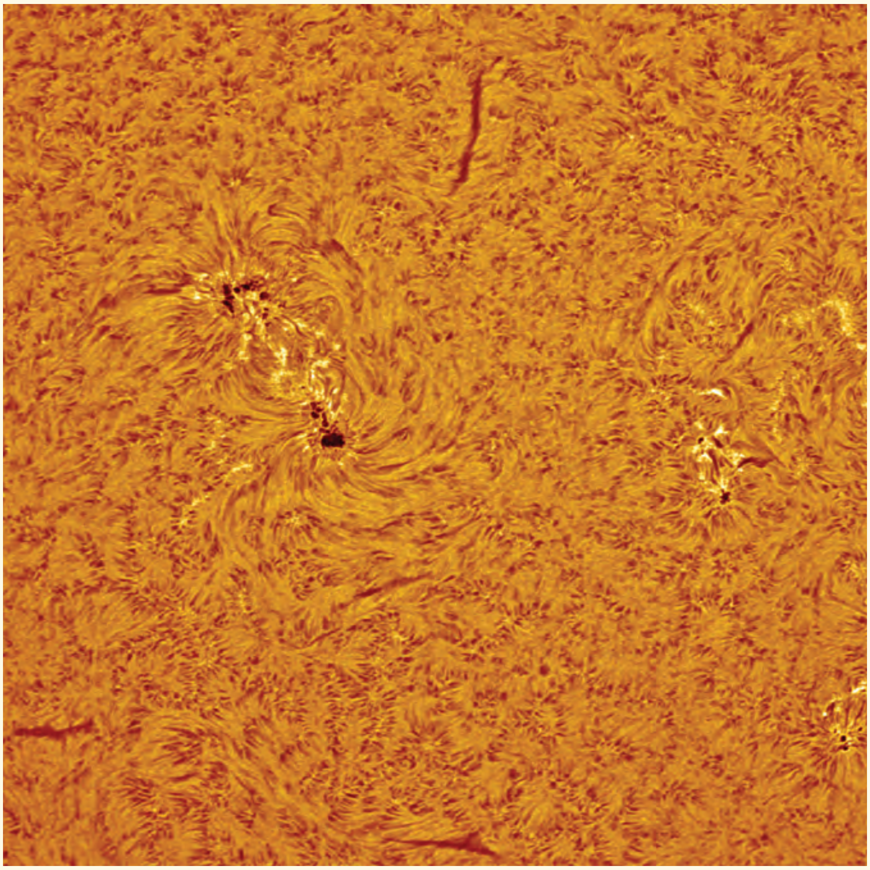
Slow solar wind traced to Sun's active regions

Multifaceted observations of the Sun reveal that interactions between magnetic field loops expel slow-moving solar wind.

Solar wind can be either fast or slow. Fast solar wind can flow more than 500 km/s, and its origins in the corona are relatively well understood. Slow solar wind is highly variable, and heliophysicists have debated where it originates in the Sun's atmosphere. By the time slow solar wind reaches Earth, it becomes difficult to resolve the slight variations in its chemical composition and thus more difficult to trace its specific solar source region. For the past two years, NASA and the European Space Agency's *Solar Orbiter* has been moving around the Sun, and it's been getting close enough to study solar wind in greater detail.

Now Stephanie Yardley of Northumbria University in the UK and colleagues have used multiple *Solar Orbiter* instruments to explore the origins of slow solar wind. Positioned about 0.5 AU away from the Sun in March 2022, the spacecraft took high-resolution images of the star's active regions. (Past analyses studied the wind at 1 AU.) It collected more data as the wind passed by the spacecraft a few days later. The short time delay meant that there was minimal loss in detail between the two measurements. Yardley and colleagues then used spectroscopic techniques to measure the composition of the wind—for instance, the iron-to-oxygen line ratio—and match it to areas of similar composition on the Sun.

Fast solar wind has already been linked to coronal holes—dark regions seen in extreme UV light in the corona. Yardley and colleagues traced the slow solar wind to an active region complex that consists of two active regions with both open and closed magnetic field loops; it is adjacent to a coronal hole with many open magnetic field loops. The proximity between a coronal hole and an active region complex provides a favorable configuration for the mechanism by which the solar wind is expelled.

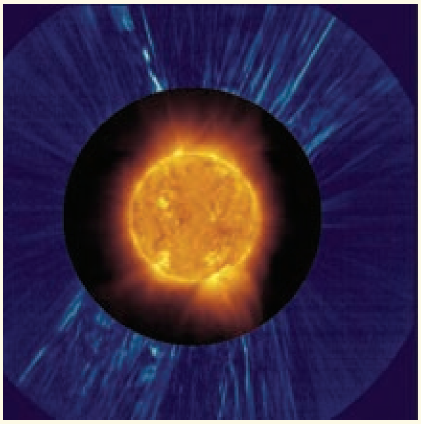


THE SUN'S SURFACE is covered in splotches: sunspots and active regions with curving filaments. The active regions are theorized to emit slow solar wind. (Image by ESA/ESAC/CESAR—A. de Burgos.)

Closed magnetic field loops in coronal regions are known to have plasma flowing along them, but nothing escapes. For plasma to be expelled from closed loops, they need to interact with adjacent open loops. The *Solar Orbiter* instruments that measure solar wind plasma and magnetic fields were able to gather evidence that the interaction between the two loop types in neighboring solar regions is what gives rise to the slow solar wind.

Yardley and her team are working on a more complex analysis for subsequent close *Solar Orbiter* approaches and on incorporating data from other missions, including the *Parker Solar Probe* (see the article by Nour E. Raouafi, Physics Today, November 2022, page 28). Future observations should deepen our understanding of what causes the variability in the wind and its origins, which will ultimately allow for better predictions of space weather. (S. L. Yardley et al., *Nat. Astron.*, 2024, doi:10.1038/s41550-024-02278-9.)

Jennifer Sieben PT



THE SUN as seen by the *Solar Orbiter* spacecraft on 25 March 2022. The inside image is of the solar surface; the outer image shows the corona. The prominent light-blue feature at the roughly eight o'clock position in the corona can be traced to an active region on the Sun's surface where magnetic field lines are interacting. (Image by ESA and NASA/*Solar Orbiter*/EUI and Metis Teams and D. Telloni et al., 2022.)

Fermilab goes deep to silence noisy radiation affecting qubits

Scientists troubleshoot to improve quantum computers.

Twin detectors at Fermilab are scheduled to start probing superconducting qubits this month as part of a growing effort to pinpoint how cosmic radiation affects the fragile components. The insights will help scientists build more robust quantum computers. The Fermilab duo consists of a new detector 100 meters underground dubbed QUIET, for the Quantum Underground Instrumentation Experimental Testbed, and one on the surface called LOUD, which began operations in 2022. Differences between the two detectors' observations will allow researchers to assess how cosmic radiation affects qubit performance.

Quantum researchers realized about four years ago that cosmic radiation limits the lifetime of superconducting qubits, says LOUD project lead Rakshya Khatiwada. When cosmic radiation interacts with a qubit, it causes decoherence, a process in which the delicate quantum state collapses and the qubit loses its stored information (see "Ionizing radiation may hinder popular qubit technology," Physics Today online, 18 September 2020). Decoherence renders qubits unusable in quantum computers.

QUIET and LOUD use superconducting qubits built with circuit loops that carry Cooper pairs—indirectly bound electrons that act as individual particles. Google, IBM, Microsoft, and other companies have chosen to build their quantum computers with superconducting qubits, and each one will need hundreds to thousands of them. Scaling up systems requires figuring out radiation's role in qubit errors, says Fermilab scientist and QUIET project lead Daniel Baxter.

Scientists set and check a qubit's state using an RF signal. At LOUD, they measure high-energy particle interactions and differentiate between radiation sources by assessing how much energy dissipates across the qubits. Muons deposit several gigaelectron volts deep in a substrate and can cause multiple qubits to malfunction, whereas beta particles deposit only a few megaelectron volts on the surface and cause localized errors. Approaches to improving qubit performance may differ depending on the culprit.

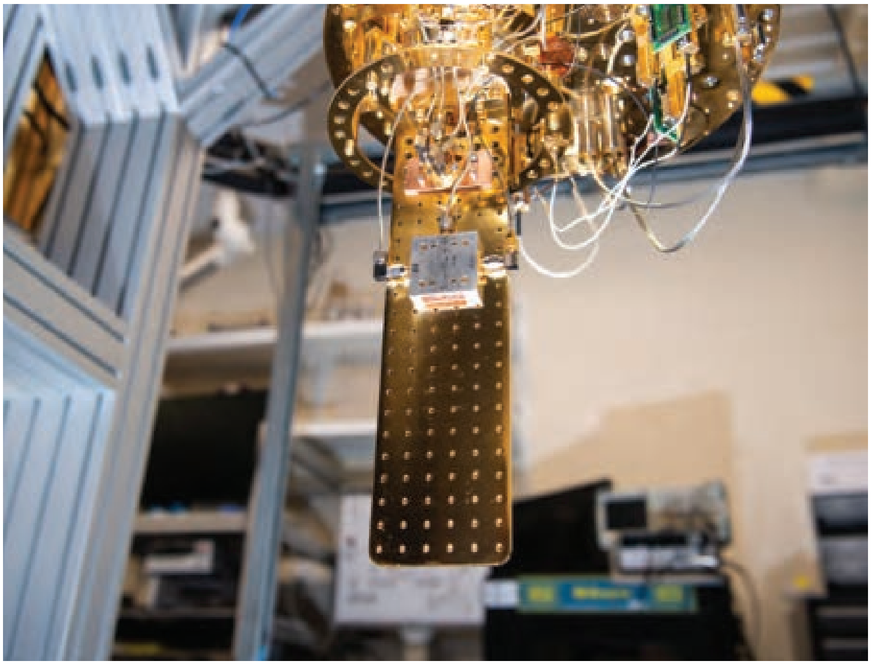
That's where QUIET comes in. Scientists will first test qubits at LOUD. They will then transfer the samples to QUIET to replicate the experiment in an environment with a 99.5% reduction in muon flux. That will allow scientists to look for energy deposits from gamma rays and other products of naturally occurring radioactive isotopes, which can disrupt qubit performance, says Laura Cardani, a researcher at Italy's National Institute for Nuclear Physics who studies supercomputing qubits at the underground Gran Sasso National Laboratory. She says that even the printed circuit board, a component of a superconducting qubit, can produce a trace amount of radiation.

DAN SVOBODA/FERMILAB

THE QUANTUM UNDERGROUND INSTRUMENTATION EXPERIMENTAL TESTBED officially opened on 30 May at Fermilab. Among the attendees at the ribbon-cutting ceremony were Dan Baxter, the lead on the experiment (third from left); Fermilab Director Lia Merminga (fourth from left); and Travis Humble (fifth from left), director of the Quantum Science Center at Oak Ridge National Laboratory, which funded QUIET.

A group effort

In May Fermilab hosted a workshop about the impacts of radiation on superconducting qubits. "It was one of the first large gatherings of people from both the quantum computing and high-energy physics communities," says Baxter. The workshop celebrated QUIET's debut and provided a space for experts across fields to discuss



ANDREA STARR/PACIFIC NORTHWEST NATIONAL LABORATORY

A SUPERCONDUCTING QUBIT (inside the gray box on the copper pegboard) is prepped to undergo testing at Pacific Northwest National Laboratory.

shared challenges in combating radiation effects in their research.

Under the 2018 National Quantum Initiative, the Department of Energy and other agencies were called on to create and fund at least two quantum research centers each. DOE invests a total of $125 million per year in five centers. One of them, the Quantum Science Center, based at Oak Ridge National Laboratory, funds QUIET. The center's director, Travis Humble, says the experiment costs on the order of $1 million, with the dilution refrigerator alone clocking in at about $500 000.

QUIET also may aid in the search for dark matter. Traditional dark-matter detection techniques are mostly limited to electron volt and higher energies, but qubit sensors could detect lower-energy dark-matter candidates. "It turns out that bad quantum computers make good quantum sensors," says Humble.

In addition to QUIET, the Northwestern Experimental Underground Site, located in the same cavern, is used for related work. Pacific Northwest National Laboratory in Washington State, the Sudbury Neutrino Observatory in Canada, Gran Sasso in Italy, and the Stawell Underground Physics Laboratory in Australia are also taking advantage of overburden shielding to investigate the effects of cosmic radiation on superconducting qubits.

At QUIET, says Baxter, electronics have been tested over the past few months to "get the noise and attenuation just right." The detector should be ready for data collection by the end of the month.

Hannah H. Means

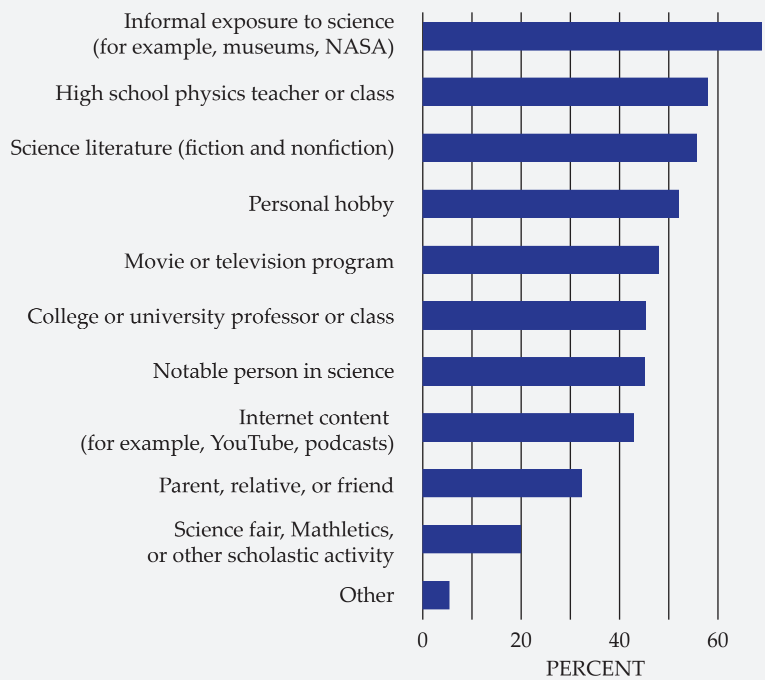
Survey asks majors: Why physics?

Informal exposure to science was cited by nearly 70% of new physics bachelor's degree recipients as influencing their choice of major. Examples include visiting science museums, partaking in science camps or programs, and attending science events. The next most frequently cited influence, selected by nearly 60% of graduates, was a high school physics teacher or class.

Those data come from the survey responses of more than 3000 people who received bachelor's degrees in physics at US institutions in the classes of 2021 and 2022. The question about influences is a newer addition to the survey of recent physics graduates that statisticians at the American Institute of Physics (publisher of Physics Today) have been conducting since 1976. For more information, see https://ww2.aip.org/statistics/physics-bachelors-influences-and-backgrounds.

Sonja Boettcher

Influences on physics bachelor's degree recipients' decision to pursue physics, classes of 2021 and 2022 combined



Source: aip.org/statistics

Optical telescopes get rigged for daytime astronomy

Applications include imaging bright stars and tracking satellites.

Astronomers want to study Betelgeuse continuously, but for a third of the year it's too close to the Sun to see in the night sky. "The development of techniques to observe during daylight fills in gaps," says Andrea Dupree, of the Center for Astrophysics | Harvard & Smithsonian, who studies the variable supergiant star.

Filling those gaps are Otmar Nickel, a retired medical physicist and astronomy hobbyist in Mainz, Germany, and Sarah Caddy, who is wrapping up her PhD at Macquarie University in Sydney, Australia. Over the past few years, they independently developed capabilities for daytime optical observation.

During daytime, seeing is hindered not only by the brightness of the Sun but also by the atmospheric turbulence created by changing temperatures. Another complication, says Lee Spitler, Caddy's thesis adviser, is that "high-flying leaves and bugs can look like satellites."

Nickel's and Caddy's setups rely on short exposure times and high acquisition rates. Taking many images improves the signal-to-noise ratio. They each imaged and made photometric measurements of Betelgeuse to test out their daytime telescopes.

Their data came in handy for Dupree. In late 2019 and early 2020, Betelgeuse underwent a historic dimming event. From a combination of space- and ground-based daytime and nighttime measurements, she and colleagues concluded that a huge chunk of Betelgeuse's surface had detached. For centuries the star had pulsated with a 400-day period. Since losing about a quarter of its surface material, says Dupree, the star's period is shorter and irregular. "It's like an unbalanced washing machine," she says. "It hasn't recovered since losing material."

With the year-round measurements made possible by daytime observations, says Dupree, astronomers "have been able to follow the recovery of the star's surface layers." That, coupled with hydrodynamic calculations, she says, led to understanding the cause of the ejection and the star's subsequent dynamics. Those studies of Betelgeuse could also shine light on how other stars eject material.

In addition, Betelgeuse is predicted to go supernova sometime in the next million years. "Humans haven't witnessed a galactic supernova in 400 years," says Andy Howell, a senior staff scientist at Las Cumbres Observatory. "The daytime observations would be invaluable. We can't afford to miss it."

In astronomy, daytime optical observing has the greatest potential for studying bright stars that appear near the plane of the solar system, Spitler says, and possibly variable stars, comets, and other transients that can't be viewed in the night sky because of their position relative to the Sun. Applications are niche, he says, "but they could still be important."

Daytime optical observations could also be useful for tracking satellites and space debris.

Fast cameras

Nickel uses a 10-inch telescope in his backyard in Mainz. Two years ago, he switched from a CCD camera to a high-speed camera with a CMOS detector. That technology, used in cell phones and other electronic cameras, has allowed him to increase his acquisition rate 50-fold to 10 images per second. An acknowledgement of his initial daytime imaging by a professional astronomer on the platform The Astronomer's Telegram, Nickel says, encouraged him to do more.

Caddy's interest in daytime imaging grew out of outreach she was doing at the local observatory. "I wanted to give kids the opportunity to look through a telescope during their visits," she says, "but the outreach was often during the day, and they couldn't see much."

Caddy was part of the team that built the Huntsman Telescope at Siding Spring Observatory. The remotely controlled telescope has 10 lenses, each 400 mm in diameter, and is optimized for nighttime observations of objects with low surface brightness.

CANON AUSTRALIA

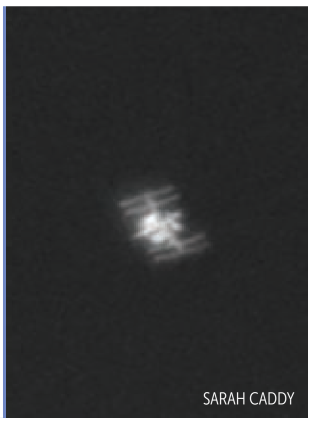
SARAH CADDY used a single 400 mm lens, like those in the Huntsman Telescope shown here, to make a daytime optical telescope. With it, she imaged and measured the brightness of the red supergiant star Betelgeuse (left inset), which is in the Orion constellation and is usually the 10th brightest nighttime star, and the International Space Station (right inset). Online, she discovered that retired medical physicist and amateur astronomer Otmar Nickel was making similar daytime observations with his backyard telescope (right).

In her pursuit of optical daytime observations, Caddy used a single such lens and a CMOS acquisition system for higher frame rates. She used high-end, off-the-shelf, and donated parts. She estimates that the total cost of her test system was less than Aus$100 000 ($67 000). "It worked," says Caddy. "We can see roughly 900 stars during the day."

Both Nickel's and Caddy's setups work best when the rough location of the

OTMAR NICKEL

SARAH CADDY



SARAH CADDY

target is known. They calibrate the telescopes at night to know where to point during the day.

Satellites and space junk

Caddy says that daytime observing could be game-changing for space situational awareness, the characterization and tracking of near-Earth objects in space. To prevent collisions, which could knock out GPS, communications, and other orbiting infrastructure, it's necessary to know where objects are in space at all times. Traditional optical tracking satellites can operate only during twilight, when the sky is dark but objects are still illuminated by the Sun. As space becomes more crowded with satellites and debris, Caddy says, daytime observations can "give us much more time to observe the rapidly increasing number of targets."

Caddy says she was shocked when she first learned about the 2009 high-speed crash between two satellites, an active commercial communications satellite and a defunct Russian one. By observing at different wavelengths during the day, researchers could obtain orientation and composition of objects in addition to position and orbit, she says, and such information could potentially inform satellites to reroute to avoid collisions.

Daytime optical telescopes may become an important complement to radar tracking facilities to "help keep space safe for the next generation," Caddy says. Companies and government agencies also monitor space objects; although they often share position information, says Spitler, they mostly keep other status details private.

Spitler says that the full Huntsman Telescope is almost ready to operate around the clock. It has CMOS detectors and filters of various wavelengths, but it still needs specialized software for daytime observations.

Toni Feder


IF/THEN® COLLECTION

Q&A: Burçin Mutlu-Pakdil chases dwarf galaxies

She navigated barriers in Turkey and the US to become an astronomy professor.

Write an essay on your ideal person. That assignment in middle school set Burçin Mutlu-Pakdil on the path to becoming a scientist. Wondering who the "cleverest person in the world" was, she started reading about Albert Einstein and his science. She became especially interested in astronomy.

That interest took Mutlu-Pakdil from Istanbul, Turkey, where she grew up, to the country's capital city of Ankara for her bachelor's degree. She earned her PhD from the University of Minnesota in 2017. An unusual galaxy she studied as a PhD student is informally named after her. Today she is an assistant professor of physics and astronomy at Dartmouth College. An observational astronomer, she studies tiny, faint dwarf galaxies; she hopes they will reveal secrets about dark matter.

In 2018 Mutlu-Pakdil was named one of the Ten Outstanding Young Persons of the World, a program run by the nonprofit, nongovernmental organization Junior Chamber International. The following year she was named an IF/THEN Ambassador, as part of a program created by the American Association for the Advancement of Science and Lyda Hill Philanthropies to inspire women to pursue STEM fields.

PT: Describe your education.

MUTLU-PAKDIL: In Turkey, universities accept you based on your performance on an exam at the end of high school. Physics doesn't require as high a score on the exam as medicine or engineering. My score was high enough to go into medicine or engineering, but I wanted to study physics. Unfortunately, when you get a physics degree in Turkey, people think you will become only a teacher, not a scientist.

I chose Bilkent University because at the time it was the best research institution and had the highest-ranking physics program in Turkey. The university is in Ankara, about a six-hour drive from Istanbul. Some people criticized my family, saying that "girls shouldn't live by themselves."

But growing up, my sister and I had always heard our father's stories about not getting an education beyond fifth grade due to financial issues. That motivated us. He kept saying, "To do good things in society, you need an education. To have a role in society, you need to get an education." So when I wanted to go to college and said I would leave my hometown, my parents were super supportive.

PT: Why did you come to the US for graduate studies?

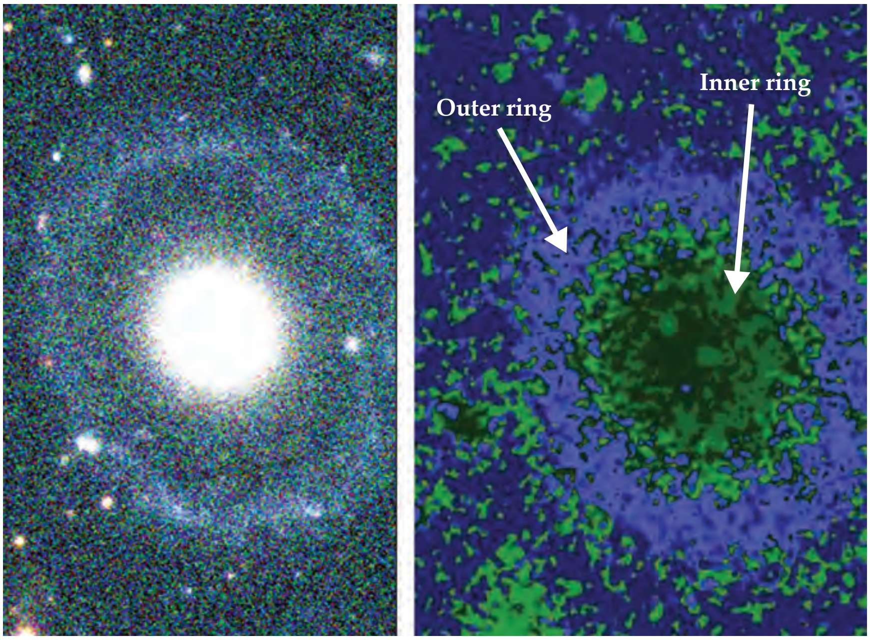
MUTLU-PAKDIL: When I went to college, I decided to practice hijab. At the time in Turkey, the law was that you cannot wear hijab in any public institution, including universities.

There were 4 women studying physics out of about 25, and I was the only one practicing hijab. I could wear hijab on campus, but not in the classroom. I used hats and turtlenecks to cover myself. It was a very awkward situation that affected my performance.

Although I enjoyed my classes, I felt like a secondary citizen in my own country. That's why, in 2009, after I graduated, I wanted to leave Turkey, not only to get a higher education but also to live my true self.

PT: You went to Texas Tech University. Why?

MUTLU-PAKDIL: I knew a person from my college who got into Texas Tech. I thought, I will not be alone. It was the only place I applied. I got in. And then I realized there was no astronomy program. I did a master's degree in experimental biophysics. I didn't like it, be-



BETTER KNOWN AS BURÇIN'S GALAXY, the elliptical galaxy PGC 1000714 has two rings. The outer one stands out in the false-color image at left; the inner ring is also visible in the color index map at right. (Left image from Ryan Beauchemin; right image from B. Mutlu Pakdil et al., *Mon. Not. R. Astron. Soc.* **466**, 355, 2017.)

cause I am a bit clumsy. Experimental labs are not my environment.

I still really wanted to do astronomy. During my master's degree, I learned how American universities work. I finally started a PhD in astronomy at the University of Minnesota.

PT: What was the focus of your studies?

MUTLU-PAKDIL: My PhD research was on scaling relationships between galaxies and their supermassive black holes. After I did my first paper, my adviser asked if I wanted to follow up on an interesting object that he and others had spotted in their survey. They thought maybe it was a second Hoag's Object—an elliptical galaxy with a blue ring around it. It's a puzzling object, and we still do not know how it formed, how it became so symmetric, or how the ring structure formed.

Later I discovered that the new object had a second ring hiding in its central bright body. We were having a hard time explaining the outer ring in Hoag's Object. Now we had an object that had an inner ring with a different color. That made it even harder to explain.

We published a paper about the galaxy, and we made a fun video about it. In the video, we called it Burçin's Galaxy. It has a catalog number—PGC 1000714—but people started calling it Burçin's Galaxy.

PT: What are you doing now?

MUTLU-PAKDIL: When I started postdoctoral research, I got really excited about other types of extreme galaxies: small, faint dark matter–dominated galaxies. Cosmological models predict that they must be the most numerous galaxies in the universe. But since they are faint and small, these dwarf galaxies are hard to find, and we have a very limited understanding of their formation.

PT: What makes those galaxies interesting?

MUTLU-PAKDIL: We know that 85% of matter in the universe is dark matter. Dwarf galaxies barely have stars; they barely have gas. They are basically the cleanest laboratories we can find to study dark matter. By counting how many dwarf galaxies are out there and determining how they are distributed in the sky, we can put significant constraints on the nature of dark matter.

PT: What did it mean to you to be named one of Ten Outstanding Young Persons of the World?

MUTLU-PAKDIL: I was named an Outstanding Young Person for Turkey after the discovery of Burçin's Galaxy. I was too busy to go to that ceremony or the ceremony in India for the world award. But my family went in Istanbul and accepted the award for me. You cannot imagine how emotional they got. Professors congratulated my father about me. Seeing my family proud and happy was the most important award I could get.

PT: How did you build up your research group at Dartmouth?

MUTLU-PAKDIL: My group has two postdocs, two graduate students, and four undergraduates. Experience can be acquired, and it's simpler to teach a technique than to inspire someone to work on something they're not passionate about. Passion is the key quality I seek in my group members.

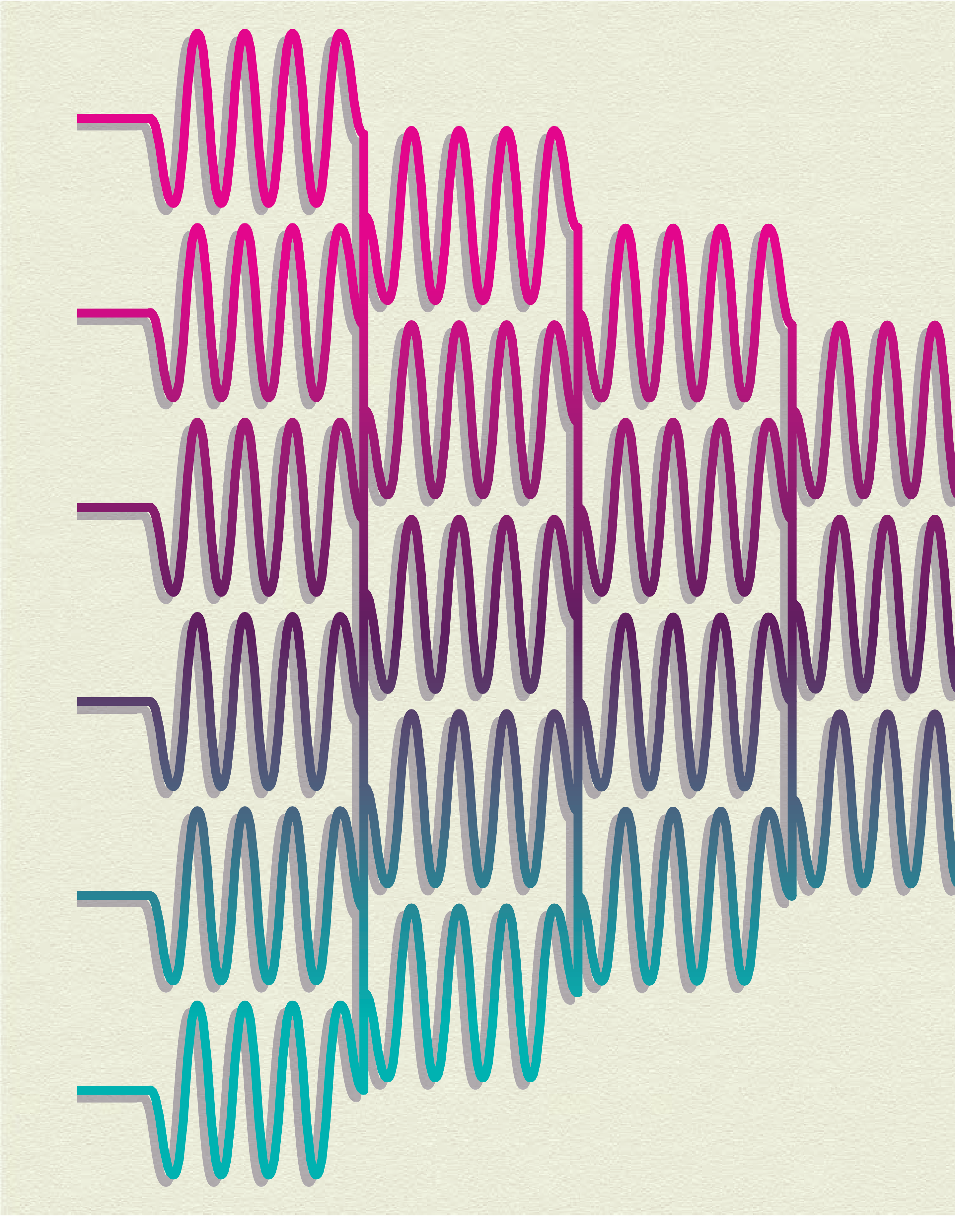
PT: Is there anything you'd like to add?

MUTLU-PAKDIL: After the discovery of Burçin's Galaxy, a friend called me and said, "You are an inspiration for our kids!" Then she said, "How about your husband? Are you reuniting with him?" I had met my husband when I was in the PhD program, and we maintained a long-distance relationship. I said, "No, we will continue long distance." Then the friend said, "You know, a career is good, but the family is more important. Maybe you should give up your career and go live with him."

I pointed out that she had just said I was an inspiration for her kids, and now she was saying I should quit. I asked my husband if people tell him that family is more important than his job, that he should quit his job and go to his wife. They don't. It's a hypocritical thing that people do only toward women, telling them to value family over career.

Once I came to Dartmouth, my husband joined me. We had thought we would unite after each step. But a postdoc is short, and he didn't find a job where I was. Unfortunately, this is very common in academia. One time, I remember as a postdoc complaining about being long distance, and then I realized that everyone in the meeting was in the same situation. It's a big systemic problem.

Toni Feder PT



Radha Boya is a professor and Royal Society University Research Fellow in the department of physics and astronomy and **Ashok Keerthi** is a Presidential Fellow in the department of chemistry at the University of Manchester in the UK. **Muhammad Sajeer Parambath** is a joint PhD scholar at the University of Manchester and the Indian Institute of Science Bengaluru.

THE WONDERLAND OF ANGSTROFLUIDICS

Radha Boya, Ashok Keerthi, and Muhammad Sajeer Parambath

Fluids confined to molecular-scale channels exhibit properties not seen on the macroscale. Research into angstrofluidics has only just begun to unlock the potential applications that the unique properties provide.

Fluids are everywhere. They are the water we drink, the air we breathe, and the blood that flows through our bodies. In nature, water striders seem to defy physics, but they simply combine surface tension of the water with the hydrophobicity of their legs to walk on water. Fluids have unique properties, including surface tension, viscosity, and capillary action, that engineers have taken advantage of to create practical applications, such as reverse osmosis, inkjet printing, and DNA sequencing.

The bulk properties of fluids are usually explained by the Navier–Stokes equations, which connect various fluid parameters, including density, pressure, and velocity. The equations describe exceptionally well the behavior of fluids at the macroscale—for example, the airflow around an airplane or the water in a kitchen tap. But the equations break down at the molecular scale.

The Navier–Stokes equations assume that the fluid is in the hydrodynamic limit, in which the continuum framework holds and the measurement scale is large enough that the tiny details of molecular movement average out. That means that the fluid's behavior can be described using a few measurable properties. The lowest scale at which the continuum framework remains valid is theoretically around 1 nm, or 10 Å. The characteristic signs of the breakdown of the continuum framework include liquid structuring and single-file transport, which lead to seemingly frictionless enhanced flows observed in angstrom-scale channels.

It would be reasonable to assume that fluids move sluggishly through angstrom-scale channels. Yet astonishingly fast flow rates of 1 billion molecules per second have been reported. Channels, pores, and other types of openings that are less than 1 nm fall within the angstrofluidics domain. Openings between 1 nm and 100 nm fall within the nanofluidics domain. At those minuscule scales, the effects of individual molecular interactions and high surface-to-volume ratios become more important.

Those effects lead to intriguing properties of fluids at molecular scales. The properties are more than just scientific curiosity; they can be adapted for practical applications. For instance, nano- and angstrom-scale channels can be used to mimic the kidney's filtering capacity in an energy- and cost-efficient manner.[1] Given that more than 2 billion people worldwide still don't have access to safe drinking water, the potential use of nanofluidics in water desalination holds immense promise to help solve real-world problems. Research into nanofluidics-assisted computing and memory devices expands the potential applications.[2,3]

The fields of nano- and angstrofluidics build on the nanofabrication expertise developed for semiconductors. Recent breakthroughs in fabrication capabilities and 2D materials have pushed angstrofluidics even further, so there has never been a better moment for anyone interested in the field to get started.

History and background

The basic building blocks of fluids are molecules that have a weaker attraction between them than the molecules observed in solids. That allows fluids to circulate freely and diffuse throughout a container, like water filling a glass or perfume spreading throughout a room. The ability of molecules to migrate from one place to another is quantified by the diffusion coefficient, whose value depends on the specific fluid, concentration, and temperature. A related term—mobility—characterizes how easily molecules can move under different stimuli.

For illustration, consider the case of a salt solution kept in two chambers connected by a small hole. A battery connected across the chambers will create an electric field in the liquid. In such a case, the electric field will drive positively charged ions to migrate through the hole in the direction of the field, and the negatively charged ions will move in the opposite direction. The ions' speed is determined by their mobility. The Nernst–Einstein relation explains the connection between the diffusion coefficient and mobility.

The dielectric constant determines how the electric field lines will be distributed. The magnitude of the dielectric constant is proportional to the ability of fluid to store electrical energy. Even though the bulk dielectric constant of water is approximately 80, confined water in angstrochannels has a suppressed dielectric constant of approximately 2, which arises from inhibited rotational motion of water molecules.[4]

Even though nanofluidics was initially explored as a scaled-down version of macrofluidics, its exceptional capabilities and properties led to its emergence as a separate field. Advances in theory, characterization, and fabrication have contributed to the growth of nanofluidics in the past 20 years. Engaging with the mysterious universe that exists at the smallest scales requires specially designed devices and tools. Nanofabrication developments adapted from semiconductor engineering enabled the creation of nanopores,[5,6] nanotubes,[7] and nanoslits.[8,9] Those three primary classes of devices, with their different channel dimensionalities, have enabled researchers to conduct ionic-transport experiments from which they have learned the properties of nanofluidics and developed related applications.

Advances in theory, characterization, and fabrication have contributed to the growth of nanofluidics in the past 20 years.

Understanding the nano- and angstrom-scale fluidic properties requires coupling measurement techniques with channel devices. The primary characterization tools used in the nanofluidics domain include electrochemical and electrical measurements, spectroscopy, gravimetry, and various microscopy techniques. Although those methods have been essential to the progress of nanofluidics research, they each have drawbacks. As a mechanical method, atomic force microscopy is limited by the influence of the tip perturbation on the liquid, and related capillary forces affect the measurement. Optical microscopy is hindered by diffraction limits, but transmission electron microscopy cannot always be used because a vacuum is needed. New types of characterization techniques are being developed to overcome those limitations.

Innovation through confinement

A ball rolling smoothly across a vast football field will follow a path that is relatively uninhibited. But what if the ball were navigating a maze? The walls would guide, block, and even divert it. That is analogous to the world of nano- and angstrofluidics, in which fluids interact intimately with the structures confining them.

At such small scales, the structures aren't passively containing the fluid; they actively influence its behavior. A maze's walls might have slopes or turns that change how the ball moves. Similarly, the surfaces of a nanochannel might be rough, charged, or embedded with specific binding sites. Those features can drastically affect the fluid's flow and interactions. For instance, a nanochannel's electrostatic charges or specialized binding sites might interact selectively with only certain molecules. And the channel's tight dimensions can act as selective gates that determine which molecules pass through and in what orientations. That isn't just a challenge—it's an opportunity. We can harness those interactions to design devices with specific functionalities, such as selective filtration or ion-charge discrimination.

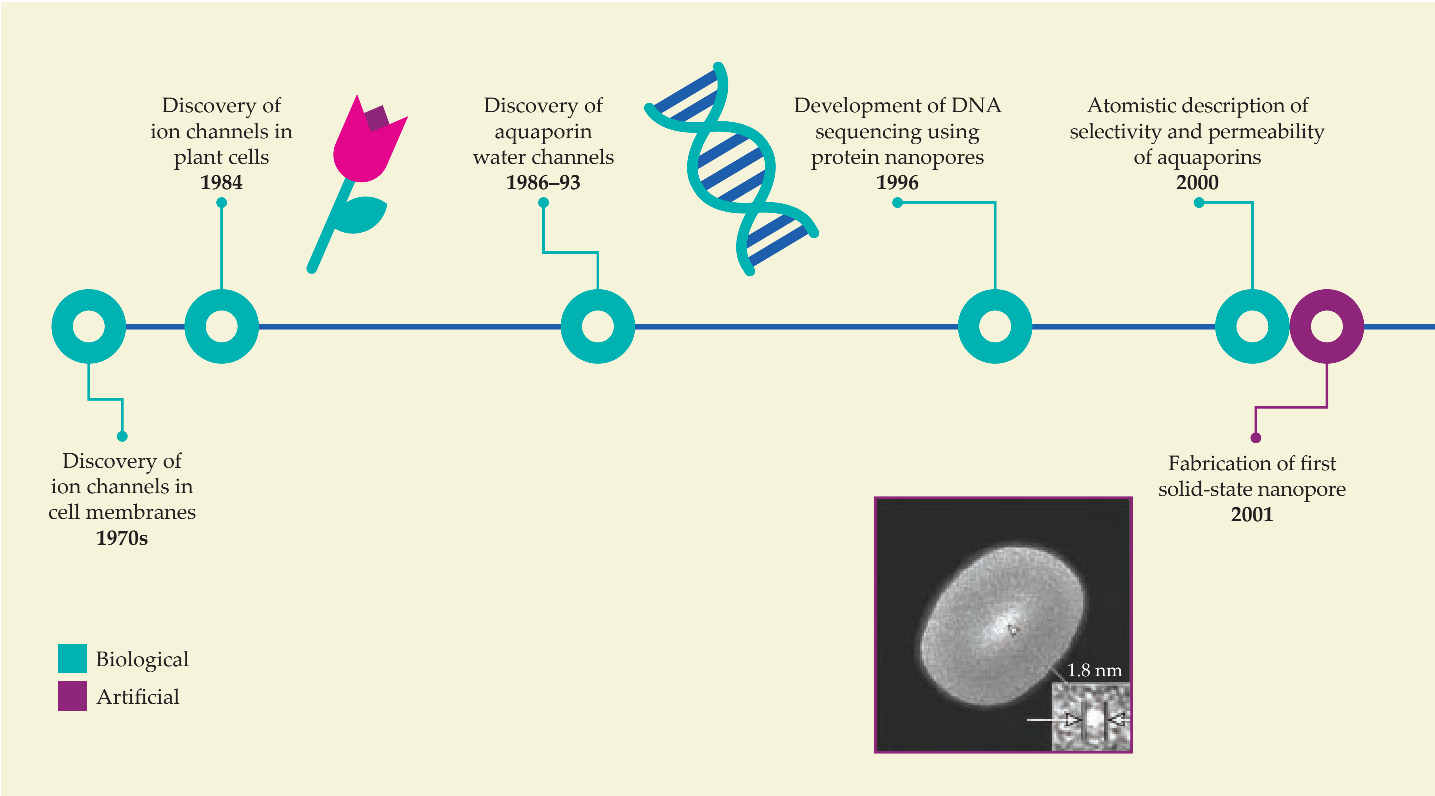
Charges on the walls of a nanofluidic channel attract ions of opposite polarity, and the accumulated ions form an electrical double layer. It is made up of the immobile Stern layer, with tightly bound ions, and the diffusion layer, with loosely bound ions. When a voltage difference is applied across a channel, the flow in the diffuse layer causes increased ion-specific mobilities in the nanochannel. The interaction between the electric double layer and the ions that are flowing through the channel results in ionic selectivity and filtration.

In angstrofluidic channels, however, wall–fluid interactions, ion-pair formation, and other phenomena lead to unique properties such as ionic streaming currents, size selectivity, and ionic memory. The new opportunities make angstrofluidics research much more breathtaking. Recent studies have examined quantum interactions between neutral molecules—for example, water confined at the nano and angstrom scales—and electronic excitations in the solid walls of channels.[10] Water flowing in a channel can induce an electronic current in the solid wall because of hydroelectronic drag, and that can be harvested as hydroelectricity from small-scale water flows. For salt solutions flowing in nano- and angstrofluidic channels, ion–ion and ion–surface interactions can lead to osmotic energy generation and neuromorphic computing applications.[2,3] Today all computers function fundamentally by controlling the electron flow through solids. Imagine a time in the future when scientists create salt-water-based computers inspired by the working of our brains!

Angstrom-scale behavior, though, isn't always different from macroscale behavior. Capillary condensation occurs across the vastly different scales. The phenomenon can be utilized in nanoporous materials to collect clean water from the atmosphere, a useful trick in deserts and other extreme environments. Another example of common behavior at the nano and angstrom scales is gas diffusion in the Knudsen regime. The Knudsen equation describes how gases move through tiny holes that are smaller than the average distance a gas molecule travels before colliding with another molecule. This equation remains accurate even for apertures that are just a few atoms wide.[9]

Real-world applications

Although angstrofluidics is a recent field of study, many applications aren't hard to imagine, especially those that build off existing nanofluidic uses. Water-filtration membranes are promising because angstrom-scale channels have



TIMELINE OF NANO- AND ANGSTROFLUIDICS. Many science disciplines overlap in the study of nano- and angstrofluidics. Tiny channels for moving ions were discovered in cell membranes in the 1970s, and later, studies of protein channels revealed their sequencing capabilities. Researchers have taken inspiration from nature and have used tools developed for other fields to create artificial channels, progressing quickly from carbon nanotubes to 2D angstrom-scale channels. (Adapted from refs. 5, 7, and 8.)

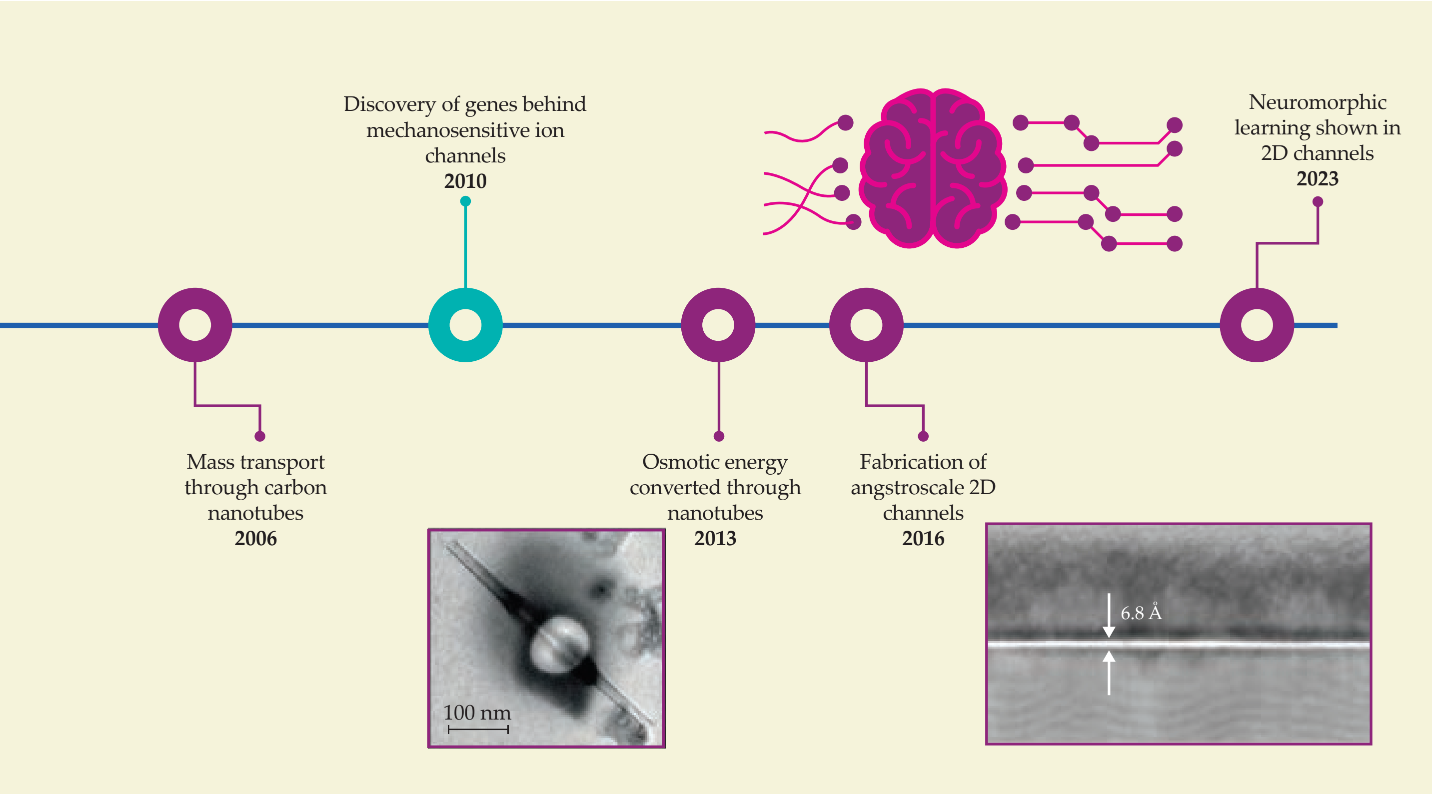
the unique ability to filter based on ionic size and charge while still maintaining high permeability.[11] Already, larger nanofluidic channels are used in desalination devices. Researchers are currently working on a practical desalination apparatus with angstrom-sized channels, but the challenge is to prevent clogging and fouling of the extremely tiny channels. Scientists also are developing efficient carbon capture and filtering technologies using nanoporous membranes.

Both nano- and angstrofluidic channels are prominent areas of research for biomedical diagnostics. Currently, nanopores are used to sense and sequence DNA and proteins for therapeutic applications.[12] Parallel efforts are underway to develop minuscule organ-on-a-chip and lab-on-a-chip devices encompassing nanochannel pathways. They have the potential to revolutionize pharmacological research by mimicking the human physiological environment on the microscale, thus enabling more accurate drug testing without the need for animal models. In drug delivery, nano- and angstrofluidic technologies are being harnessed to engineer controlled-release mechanisms and improve the delivery efficiency of therapeutic agents, especially in targeting hard-to-reach cancerous cells. Not only are nano- and angstrofluidic technologies more compact and efficient, but they also have the potential to be more cost-effective, which makes them an especially valuable tool in democratizing health care.

Remote areas can also benefit from energy storage and generation via nanochannels. Ion-selective nanochannels can be placed between solutions of different salt gradients to allow the continuous movement of specific ions. That leads to constant generation of an ionic current. Although not in use commercially at a large scale, such so-called blue-energy generation has been demonstrated at the intersections of seas and rivers. Membranes with angstrofluidic channels will have improved selectivity and thus will lead to efficient osmotic energy harvesting. Angstrofluidics may also offer a way to improve the material and design of batteries, a research goal that covers many disciplines.

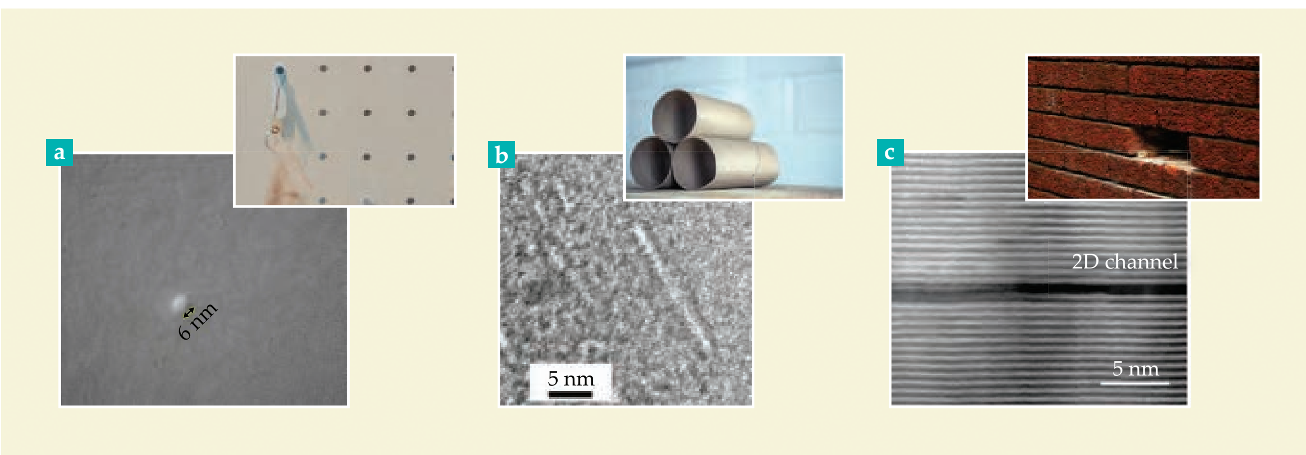
Current bottlenecks

Before we can fully use nanofluidic and angstrofluidic channels in large-scale industrial deployment, certain obstacles need to be addressed. Some of the primary bottlenecks are in imaging techniques, fabrication, and the general knowledge gap.



To reveal the mysteries at the molecular scale, new measurement techniques capable of capturing angstrom-scale dynamics are necessary. Several characterization tools are currently being explored, including simple patch-clamp amplifiers to measure ionic currents and high-end microscopes. Yet, as previously discussed, those techniques have drawbacks and limitations. New devices, with better localization than is currently possible, are needed to detect changes happening within subnanosecond intervals.

Tools to explore the interfaces between two phases of matter—liquid and the interacting solid surface—are also necessary. New and old techniques will need to be combined and correlated to allow multifaceted simultaneous measurements and imaging. And because all the phenomena explored here are at the molecular scale, noninvasive measurement techniques need to be developed.



NANOFLUIDIC DEVICES and their macroscale analogues. **(a)** A nanopore in a silicon nitride membrane[6] is similar to a hole in a peg board. **(b)** A carbon nanotube[13] is similar to a toilet paper roll. **(c)** An angstrom-scale slit in stacked molybdenum disulfide layers[9] is similar to a missing brick in a wall. (Peg board courtesy of Toa Heftiba/Unsplash; toilet paper rolls, Jessica Lewis/thepaintedsquare/Pexels; brick wall, Him_83/CC BY-NC-SA 2.0/Flickr.)

BLUE-ENERGY GENERATION, or osmotic power, is one of many potential real-world applications for nano- and angstrom-scale fluidics. Ion-selective nanochannels can be placed at the intersections of seas and rivers to generate a constant ionic current—a clean energy source. As with the use of nanoporous materials to collect clean water from the atmosphere in extreme environments, the technology has been demonstrated and is waiting on fabrication advancements to be put to use on a wider scale.

Even though nano- and angstrom-scale channels have the potential to be used in several real-life applications, there are currently no methods to scale up the fabrication of nanochannels for wide use without affecting device functionalities. For example, an industry-scale deployment of nanofluidic filters with molecular selectivity and high permeation to extract uranium and lithium from seawater would be highly beneficial. It requires thousands of device arrays, however, and with the current manufacturing capability, it is a nearly impossible task. Scalable fabrication techniques for nanochannels with anticlogging and antifouling properties are needed. Current developments, such as rapid fabrication techniques that use membrane protein nanosheets, are a step in that direction.

In biology, aquaporin—a type of protein channel—can allow one water molecule through at a time and still achieve transportation of a billion molecules in a second. The human kidney has excellent filtration. Ionic pumps in neurons are extraordinarily fast and sensitive. Creating angstrochannels with comparable capabilities is still a distant dream.

Biological channels also are extraordinarily multifunctional. In addition to offering ion selectivity, they respond to different stimuli, control flow rate, and are highly adaptable. In contrast, artificial nanochannels have to date focused only on specific tasks. Chemists, electrical engineers, and materials scientists working together can provide the interdisciplinary expertise needed to develop nanochannels that mimic the capability of biological pores and go beyond single-purpose devices.

The importance of working with what you have

Because nano- and angstrofluidic devices are at the tiniest scales and involve several key elements—for example, force fields and complex ion-wall interactions—many aspects are yet to be fully understood. For example, the liquid molecules flowing closest to the channel surface are classically expected to have zero velocity, but that may not always be true in angstrofluidic channels. A supersmooth channel wall instead leads to an increase in the fluid velocity, but researchers do not yet have a complete understanding of the mechanisms behind it.

Studies in carbon nanotubes have observed water freezing above 100 °C. And the effect of channel defects on fluid transport is puzzling. Using existing tools and devices, researchers are working on the fringes of what's possible to decipher the secrets of nano- and angstrofluidics. It is not always the most efficient way to fill in knowledge gaps, but persistence is what keeps the field moving forward.

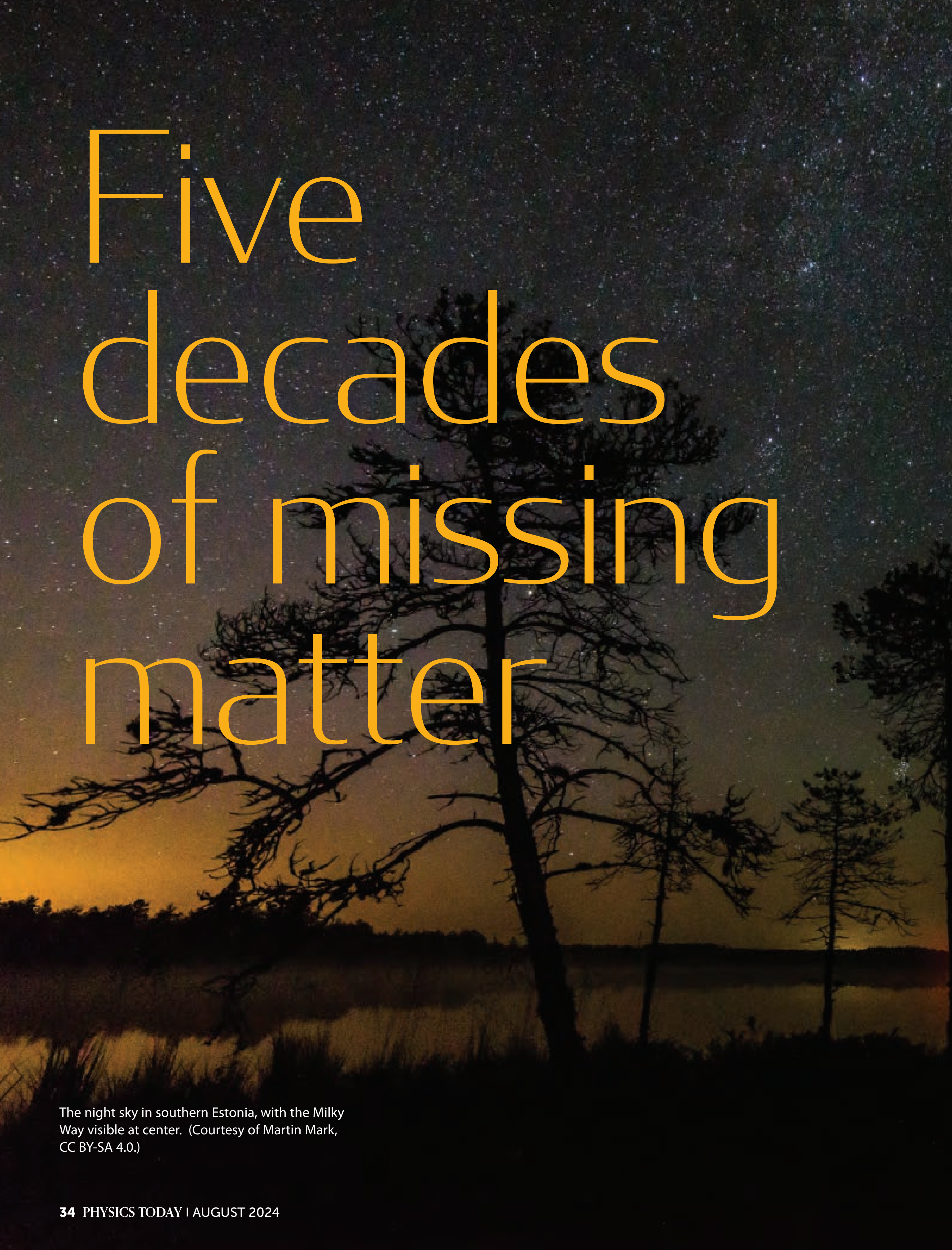
Even after 20 years of nanofluidics research, the field continues to surprise as groundbreaking discoveries reveal extraordinary properties, especially with angstrom-scale channels. As William Shakespeare wrote in *Hamlet*, "There are more things in heaven and earth . . . than are dreamt of in your philosophy." Likewise, there are undoubtedly more phenomena to be understood in the angstrofluidics domain that researchers have yet to imagine. The future of the field mostly resides in technological developments and fabrication advancements. The applications span disciplines, and collaboration across the sciences will yield many benefits. Progress will require many more curious researchers ready to delve into the field and join the exciting search to unravel the mysteries of nano- and angstrofluidics.

This work was supported by the Royal Society University Research Fellowship, a European Union H2020 ERC Starting Grant (852674—AngstroCAP), the Philip Leverhulme Prize (PLP-2021-262), an EPSRC New Horizons grant (EP/V048112/1), and a Commonwealth Split-Site PhD Scholarship. We are grateful to everyone who contributed to the research on angstrom-scale fluidics.

REFERENCES

1. S. Marbach, L. Bocquet, *Phys. Rev. X* **6**, 031008 (2016).
2. P. Robin et al., *Science* **379**, 161 (2023).
3. T. Emmerich et al., *Nat. Electron.* **7**, 271 (2024).
4. L. Fumagalli et al., *Science* **360**, 1339 (2018).
5. J. Li et al., *Nature* **412**, 166 (2001).
6. A. Dominic et al., *J. Vac. Sci. Technol. B* **41**, 043204 (2023).
7. A. Siria et al., *Nature* **494**, 455 (2013).
8. B. Radha et al., *Nature* **538**, 222 (2016).
9. A. Keerthi et al., *Nature* **558**, 420 (2018).
10. N. Kavokine, M.-L. Bocquet, L. Bocquet, *Nature* **602**, 84 (2022).
11. A. Esfandiar et al., *Science* **358**, 511 (2017).
12. H. Brinkerhoff et al., *Science* **374**, 1509 (2021).
13. R. H. Tunuguntla et al., *Nat. Protoc.* **11**, 2029 (2016).

Five decades of missing matter



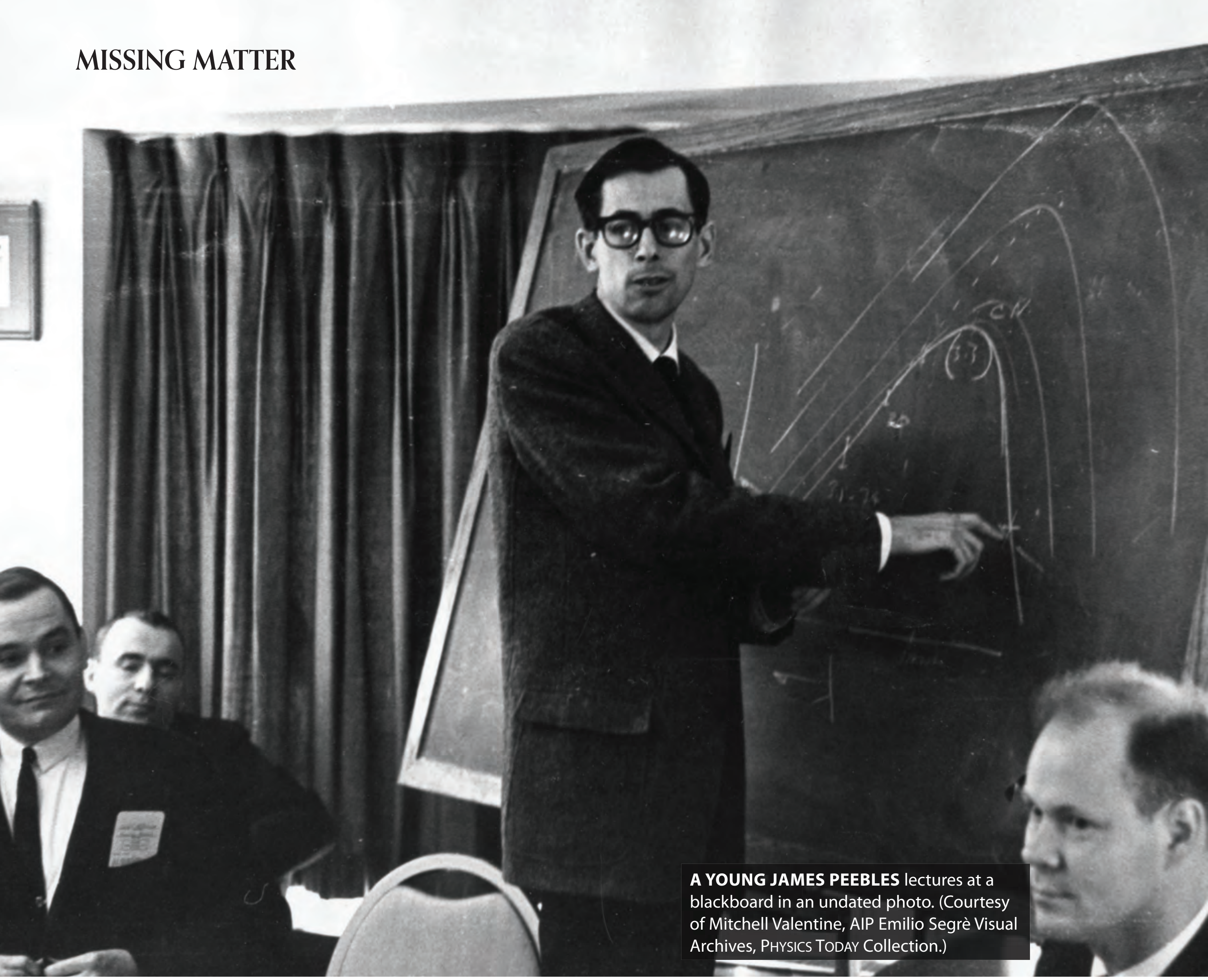
The night sky in southern Estonia, with the Milky Way visible at center. (Courtesy of Martin Mark, CC BY-SA 4.0.)

Jaco de Swart is an AIP Robert H. G. Helleman Memorial Postdoctoral Fellow at MIT in Cambridge, Massachusetts. A trained physicist and historian, he is currently completing a book manuscript on the history of dark matter.

Jaco de Swart

Groups on both sides of the Iron Curtain hypothesized in the 1970s that a yet-unknown material makes up most of the mass in the universe.

Two papers that appeared in 1974 changed the face of the universe. Independently authored by separate collaborations, one in the US and the other in Estonia, they argued that galaxies are 10 times as massive and extensive than had previously been thought. Both groups combined various astronomical observations to show that most of the universe's mass is hidden in invisible clouds around galaxies. The universe itself, too, they illustrated, is heavier by a factor of 10 than had previously been believed, potentially changing human understanding of the fate of the cosmos. Their arguments marked "a watershed in our understanding of galactic structure, galaxy formation, and cosmology," read a review in the 1999 centennial issue of the *Astrophysical Journal*.[1] Five decades ago those papers proposed the existence of what we now know as dark matter.



A YOUNG JAMES PEEBLES lectures at a blackboard in an undated photo. (Courtesy of Mitchell Valentine, AIP Emilio Segrè Visual Archives, PHYSICS TODAY Collection.)

Today dark matter is not only one of the pillars of modern cosmology but also one of its central conundrums. The existence of unseen matter distributed throughout the universe is key to understanding cosmic structure and evolution: It explains how galaxies move about and why they exist in the first place. But at the same time, after decades of dedicated research and experimentation, the exact nature of dark matter—what the stuff is actually made of—is still unknown. Currently several dozen massive international experimental efforts, including ones in underground mines and in space stations, are attempting to detect evidence of hypothesized dark-matter candidates. The two papers from 1974 formed the basis of that profound hypothesis and initiated an exhilarating new era in cosmic understanding.

Here I tell the story of how those two papers made dark matter come to, well, matter. That story is unlike usual dark-matter histories, which typically center on the roles of astronomers Fritz Zwicky and Vera Rubin. In the 1930s Zwicky found that galaxies in clusters are unstable without extra mass, and in the 1970s Rubin observed that galaxies rotate faster than their luminous mass would imply. Astronomy textbooks normally cite those observations as evidence for the existence of dark matter.

But facts and observations themselves do not tell a history (see the article by Matt Stanley, PHYSICS TODAY, July 2016, page 38). To understand the origin of the case for dark matter, we need to know how prior observations made by Zwicky, Rubin, and others were interpreted to be evidence for its existence. In what context were they used to show that the universe had preponderous amounts of missing matter? Who started to care, and why? That happened independently 50 years ago on both sides of the Iron Curtain.

A search for two numbers?

Half the story starts with a prolific young astrophysicist named Jeremiah Ostriker. An expert on stars, Ostriker received a BA in physics and chemistry from Harvard University in 1959 before pursuing a PhD at the University of Chicago with Subrahmanyan Chandrasekhar, who was famed for groundbreaking work on stellar astrophysics that would earn him the 1983 Nobel Prize in Physics (see the article by Freeman Dyson, Physics Today, December 2010, page 44). Under Chandrasekhar, Ostriker started a career in the physics of stars and their rotation. His PhD research was devoted to showing that there is a hard limit to how fast stars can rotate before they disintegrate. After a brief stint at the University of Cambridge, Ostriker received an assistant professorship in 1965 at Princeton University, where he continued his influential work on stellar physics.[2]

Ostriker's area of research—the properties and evolution of stars—had dominated astronomy since the 1930s. But the field's focal point began rapidly shifting in the 1960s. With the aid of Cold War–era technological developments, astronomers began opening new windows to the universe with observations across the electromagnetic spectrum. Interest resurged in Einstein's theory of general relativity. The astronomical workforce increased dramatically, and a new generation of researchers began observing novel phenomena on the galactic scale and beyond, including quasars, pulsars, and the cosmic microwave background. Cosmology became a new focus for young astrophysicists.[3]

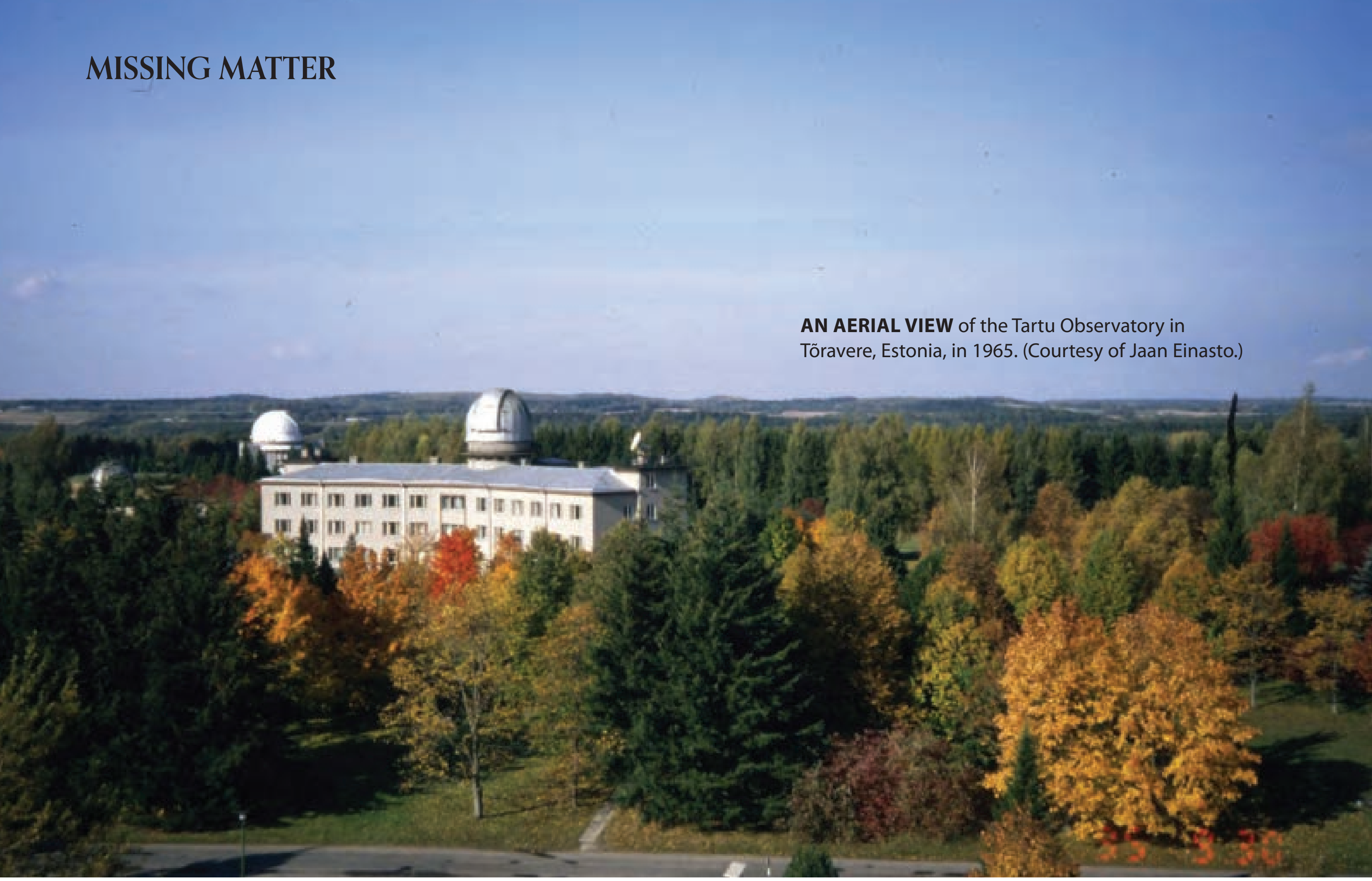
That new generation, Ostriker included, began to part ways with Edwin Hubble's classical cosmological endeavor, which was termed in the headline of a February 1970 Physics Today article by Allan Sandage "a search for two numbers": the Hubble constant, measuring the expansion of the universe, and the deceleration parameter, quantifying the rate at which the expansion is slowing. In the late 1960s New Zealand astronomer Beatrice Tinsley and colleagues had shown that the brightness of galaxies changes as they age. As Ostriker recalled in an interview, that work aroused "suspicion" of the traditional enterprise. "All of a sudden we realized galaxies have to evolve," he said.[4] With that revelation, some of Hubble's classic cosmological tests were deemed unreliable. Researchers instead began using observations to work out the many physical processes that govern galaxies and the universe.

By 1971, when Ostriker was promoted to full professor at Princeton, he had shifted his focus from stars to galaxies. He used his expertise in the evolution of stars to show how stellar processes could influence the total luminosity of galaxies during their lifetime. Another question Ostriker delved into was inspired by his graduate work on rotating stars: How do rotating galaxies maintain stability during their lifetime? Answering that question, however, required modeling galaxies with computers—a practice that Ostriker was not familiar with. He turned to his colleague James Peebles for help.

Can galaxies survive?

Born and raised in Winnipeg, Manitoba, Canada, Peebles got a BS from the University of Manitoba before moving to Princeton to study physics in 1958. Although he started in particle physics, Peebles eventually became charmed by the work of Robert Dicke, a gravitational physicist who by 1960 was well known for his unique approach, which involved

JEREMIAH OSTRIKER speaking in China in 1980. (Courtesy of the AIP Emilio Segrè Visual Archives, gift of Jeremiah Ostriker.)



AN AERIAL VIEW of the Tartu Observatory in Tõravere, Estonia, in 1965. (Courtesy of Jaan Einasto.)

experimentally testing different gravitational theories.[5] Dicke asked Peebles in 1964 to consider the consequences of a potential remnant from the universe's hot and dense early history: an observable cosmic background of microwaves in the sky. When the cosmic microwave background was observed in 1965, Peebles was standing at the cradle of the modern Big Bang theory. One of the pioneers of a new, "physical" cosmology, he worked on such topics as the synthesis of nuclear elements in the Big Bang and the formation of galaxies and the cosmic structure—work for which he received a share of the 2019 Nobel Prize in Physics.

When Ostriker knocked on his door, Peebles had just come back from a visit to the famed nuclear research facility in Los Alamos, New Mexico, where he was invited to help make sense of a highly energetic flash of gamma radiation detected by a satellite in 1969. (Although detected by a satellite meant to monitor nuclear weapons tests, that flash was later recognized as the first detection of a gamma-ray burst.) Peebles made good use of his visit: He used the facility's supercomputers—normally used to model nuclear weapons and explosions—to create the first simulations of galaxy clustering in the universe. It showed how a homogenous soup of mass would increasingly grow clumpy under gravity.

Following Ostriker's suggestion when he arrived back in Princeton, Peebles used his computer-punch-card dexterity and "banged out some n-bodies."[6] To be precise, it was 500 bodies: His model simulated the stability of galaxies by using starlike particles moving under gravity in a disk. Ostriker and Peebles quickly found that rotating galaxies in their model were "rapidly and grossly unstable." In their computation, disks of stars disintegrated after a single rotation.[7]

Something was off. It was well known that the Milky Way had existed for much longer than a single rotation. So how did it survive in the real universe? One explanation, they argued, involved rethinking the distribution of galactic mass. Instead of positing that all the mass was located in the bright disk, they proposed that more mass was in the spherical bulge of the galaxy, which they termed a "halo." That halo would help stabilize a galaxy as it rotated.

Few galactic dynamicists were enthusiastic about Ostriker and Peebles's new idea. Among other reasons for skepticism, it was unclear if their analysis would hold for all galaxies. Would it account for galaxies that are not rotationally symmetrical? If so, is a massive halo the only way to prevent instability? Many astronomers agreed with MIT astronomer Alar Toomre, who called the idea "a real migraine" a few months after Ostriker and Peebles's paper came out in 1973.[8]

The mass of the universe

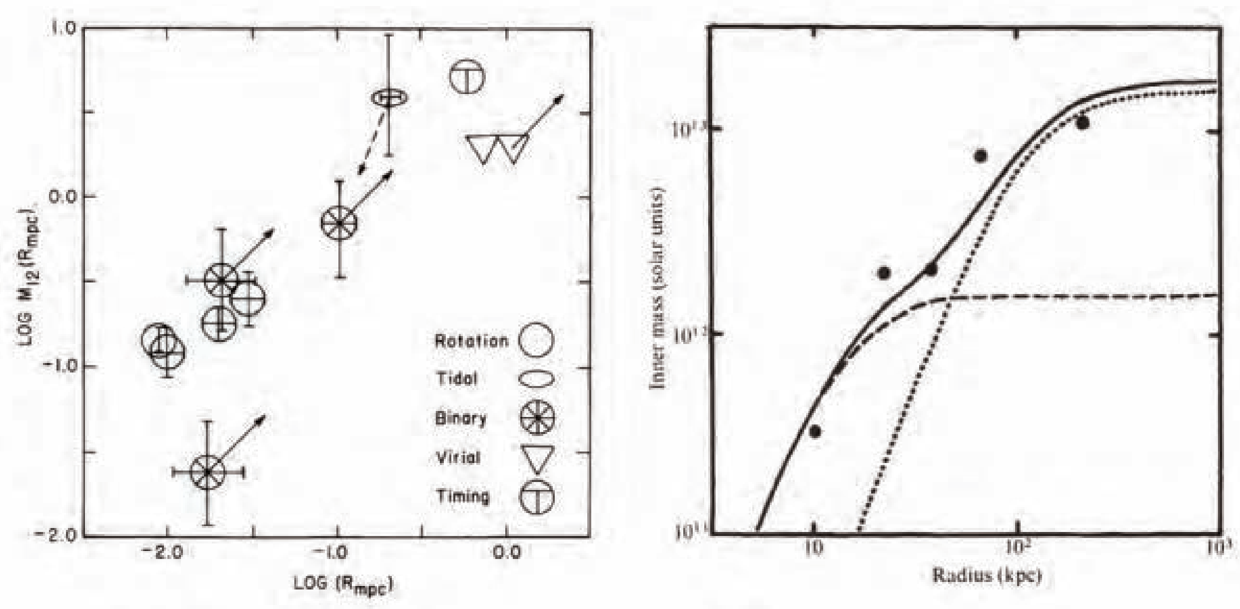
Although Ostriker and Peebles's stability argument remained controversial, it managed to inspire Amos Yahil, a lecturer at Tel Aviv University in Israel, on leave at the time. Unsatisfied with the field of particle physics—in which he obtained his PhD at Caltech in 1970—Yahil had recently shifted to cosmology. While a postdoc at the Institute for Advanced Study in Princeton, New Jersey, in 1971–73, he began studying the distribution of mass in the universe, which was understood to be one of the key parameters to understanding the fate of an expanding cosmos. Was the universe "open," meaning that it would expand forever? Or was it "closed," meaning that it would have so much mass that gravity would cause it to collapse together again sometime far in the future?

To study the universe's mass distribution, Yahil began modeling clusters of galaxies. Although other researchers had modeled galaxy clusters by using clear spatial boundaries, Yahil observed something curious: Galaxy counts suggested that there was no clear end to a cluster. In other words, a cluster's density appeared to have no sharp spatial cutoff. After hearing Ostriker talk in 1972 about the possibility of massive galactic halos, Yahil started to wonder whether galaxies, like clusters, also lacked a clear point where their mass suddenly ends. He sent a draft manuscript on the subject to Ostriker, initiating their fruitful collaboration.

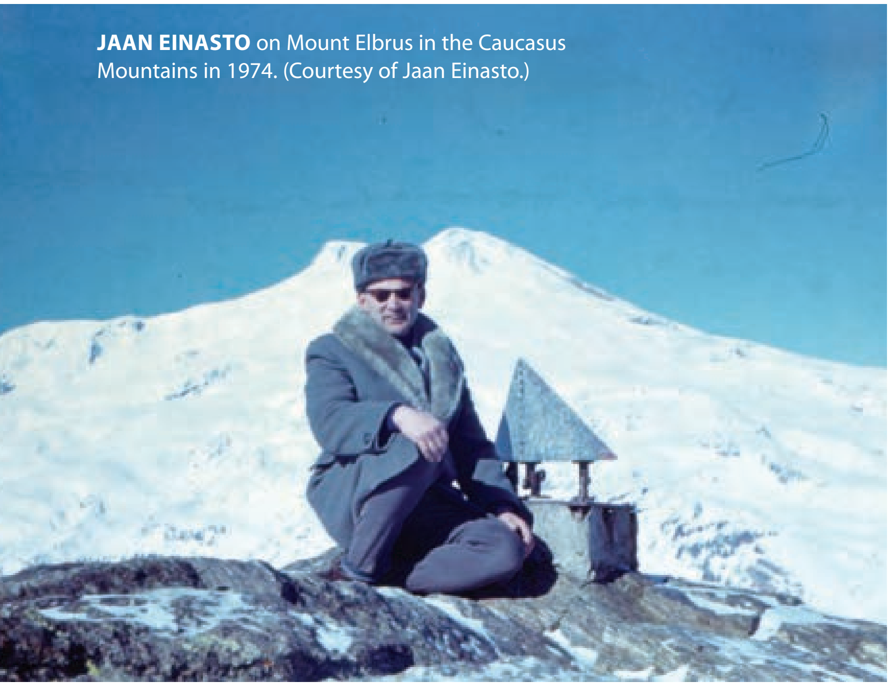
Ostriker himself had been similarly inspired by the consequences of his halo idea. "If the disk doesn't dominate the [mass of a galaxy's] interior, then maybe it doesn't dominate the exterior either," he recounted during an interview. Ostriker and Yahil began collaborating to see whether their idea indeed held. What if a spherical dark component increasingly dominates the mass of a galaxy in its outer edges? The duo soon asked Peebles, who had just written a textbook on cosmology and the mass distribution in the universe, to join their collaboration.

The team gathered dynamical measurements of galaxy masses from various astronomical subdisciplines. Those determinations were based on the gravity needed to explain the movement of cosmic objects. They included galaxy pairs, small groups and large clusters of galaxies, and rotating galaxies. The three men found that those dynamical measures of galaxy mass kept increasing with a galaxy's radius: In other words, the farther out the mass of a galaxy is measured, the higher the mass of the system—even if it was measured outside of the luminous disk of the galaxy. "The masses of ordinary galaxies may have been underestimated by a factor of 10 or more," they concluded in their paper, which was published in October 1974.[9] Galaxies had no clear boundary but were surrounded by extended, invisible massive halos, possibly of "faint stars." Added together, those hidden halos

Demonstrating the existence of dark matter



Diagrams from the landmark 1974 papers by the Princeton University and University of Tartu groups that demonstrate the existence of large halos of unseen mass surrounding galaxies. The diagrams plot different masses of galaxies as measured within a distance *R* from their centers. They show how the mass of a galaxy does not stop at a fixed point but instead keeps increasing linearly with the radius far beyond the bright visible disk of a galaxy. At left is a plot from the article by Jeremiah Ostriker, James Peebles, and Amos Yahil. The data points signify observed mass as determined by different methods, including ones based on galaxy rotation, galaxy pairs, and cluster dynamics (labeled "virial" in the diagram after the statistical mechanics theorem used to determine the total mass of galaxies in a cluster). At right is a plot from the article by Jaan Einasto, Ants Kaasik, and Enn Saar. The dots represent the observed values obtained from five groups of galaxy pairs, the dashed line is the mass function of known stellar populations, the dotted line is the implied mass distribution of the "dark" corona, and the solid line is the total mass distribution of the galaxy, including the corona. Mass is given in units of solar mass $M_{\odot}$. (Left diagram from ref. 9; right diagram from ref. 15, *Nature* citation.)



JAAN EINASTO on Mount Elbrus in the Caucasus Mountains in 1974. (Courtesy of Jaan Einasto.)

significantly increased the mass of the universe, which suggested that the universe might be closed.

The Soviet parallel

On the other side of the Iron Curtain, almost 6500 kilometers east of Princeton, a similar conclusion was drawn. As Yuri Gagarin launched into space in 1961, on land the Soviet Union was developing a strong workforce in cosmological physics. One of its main drivers was Yakov Zeldovich, a physicist who was famous for his work on the Soviet nuclear bomb project in the 1940s. He began gathering a group of bright physicists to tackle the problems of the cosmos. Working in parallel to Dicke's group in the US, Zeldovich and his team quickly became internationally renowned for their work on neutrinos, quasars, black holes, and the cosmic microwave background. Zeldovich's weekly two-hour seminar at Moscow State University's Sternberg Astronomical Institute became a central hub for anything cosmology and drew scientists from all around the Soviet Union.[10]

In 1971 one of the Sternberg seminars was given by a 42-year-old astronomer named Jaan Einasto, who had traveled by train from Tõravere, a small town 20 kilometers outside of the city of Tartu in what was then the Estonian Soviet Socialist Republic. Tõravere housed the Tartu Observatory, which Einasto had helped set up when it was relocated outside the city in 1964. Estonia's astronomical heritage dates back to the early 19th century, when the famed Baltic German astronomer Friedrich Georg Wilhelm von Struve helped make a name for the Tartu Observatory with his observations of double stars. Einasto was part of a new postwar generation in Estonian astronomy.[11]

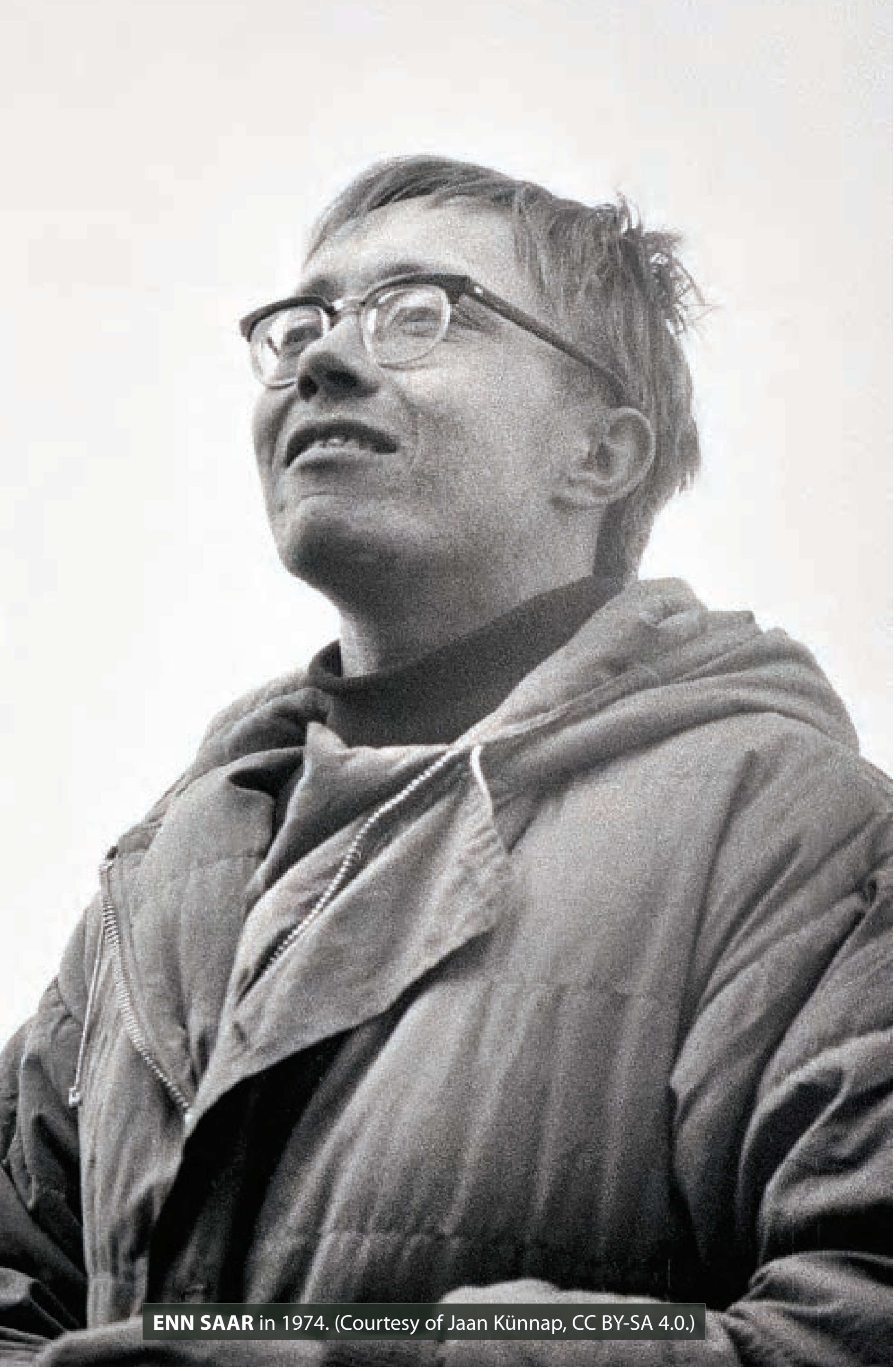
Einasto was invited to Moscow to discuss new theoretical models of galaxies—a subject he had worked on since his PhD work under Grigori Kuzmin. Working in parallel to Tinsley and other astronomers in the US, Einasto aimed to mathematically describe and understand the evolution of galaxies by modeling their luminosity and distribution of matter. Above all, he aimed to precisely model galaxies by using known populations of stars. Taking existing observations, he modeled how star populations were distributed in galactic components, such as the bulge, the core, and the disk. In Moscow, he gave a seminar on his recent model of the Andromeda galaxy, the nearest neighbor to the Milky Way. It piqued the interest of Zeldovich, who invited Einasto to present at the annual Soviet winter school for astrophysics in the Caucasus Mountains.

By 1972, after many conversations at the winter school about a model of Andromeda, Einasto had hit what he recalled in an interview as a "deadlock." He had come across a paper from Morton Roberts, a radio astronomer at the National Radio Astronomy Observatory in Green Bank, West Virginia, that described using hydrogen clouds to model Andromeda's rotation. Roberts's data showed that those clouds were moving remarkably fast along the galactic edges—beyond the visible stars in the disk. What could be out there at the edges of the galaxy that explained this? "No combination of stellar populations was able to explain rotation data of galaxies," Einasto wrote in his autobiography.[12] It was a problem Einasto was unable to address properly until he came to speak about it with a colleague at the observatory.

A bomb in the Caucasus

It was coworker Enn Saar, a cosmologist, who helped Einasto by teaching him a valuable lesson in gravitational physics. The son of two Estonian fishers, Saar grew up walking to high school with matches and a newspaper to fend off wolves—and having a deep interest in the cosmos. Shortly after defending his PhD on Big Bang cosmology in 1972 at the University of Tartu under the supervision of cosmologist Arved Sapar, Saar began discussing with Einasto the problems in modeling the outer edges of galaxies. There was "sort of a difference in attitudes," Saar said in an interview. His experience in cosmology meant that he saw little problem with extended galaxies. He said that lacking borders was "a normal state of any gravitational body." Perhaps there was more than met the eye and galaxies could extend far beyond their luminous disk.

For Einasto, Saar's insight meant "abandoning the idea that only known stellar populations exist in galaxies." They might be surrounded by a new population of yet-unknown nature. Einasto and Saar named that invisible population the


ENN SAAR in 1974. (Courtesy of Jaan Künnap, CC BY-SA 4.0.)

"galactic corona." Einasto presented his ideas at the First European Astronomical Meeting in Athens in 1972. "Giant galaxies may be surrounded by massive coronae of very large dimensions," his abstract read. Einasto suggested that that "unknown matter" might exist in the form of "rarefied ionized gas."[13] The response to his talk in Athens, Einasto recounted in an interview, was lukewarm at best. The half-hearted reaction made him determined to find more data to support his claim.

That was no small task in the Soviet Union. Iosif Shklovsky, the eminent Russian astronomer famous for his collaboration with Carl Sagan, is said to have once joked that all Soviet astronomical observations were done through the US-based *Astrophysical Journal* because Soviet astronomers were often hampered by weather conditions and equipment problems. But even obtaining issues of US journals in the Soviet Union—especially in Estonia, far from scientific centers such as Leningrad or Moscow—was not trivial. Einasto was often forced to use foreign travel stipends to acquire his own copies of publications such as the *Astrophysical Journal*.

Through an arduous search of the international literature, Einasto discovered the long-standing "Zwicky problem": that galaxies in groups and clusters seem to move so fast that they must either be exploding or require large amounts of extra mass to stabilize them. Einasto found data on groups and pairs of galaxies that would complement Roberts's galaxy-rotation data and make a strong case that coronae of unseen mass must exist. He also learned about x-ray studies that showed that galaxies did not have enough ionized gas to account for all the mass in their coronae. That knowledge prompted him to hypothesize that the mass might be made up of something akin to a new population of stars. Einasto diligently worked out the calculations with Saar and local student Ants Kaasik.

They presented their work in January 1974 at Zeldovich's annual winter school in the Caucasus. Against the backdrop of Mount Elbrus, the tallest mountain in Europe, Einasto shared his idea about galactic coronae: There must be a still-unknown nonstellar population of stuff surrounding galaxies. As he wrote later, it was "as if a bomb had exploded."[14] The avid young physicists in Zeldovich's group immediately started to do back-of-the-envelope calculations: Could coronae consist of gas or neutrino clouds? As Einasto, Saar, and Kaasik began writing up their conclusions for a Soviet astronomical leaflet, the *Astronomicheskii Tsirkulyar* (*Astronomical Circular*), their host intervened: "Zeldovich insisted this must be published in some really important journal," Einasto told me in an interview.

On Zeldovich's advice, Einasto and his group decided to translate their paper into English and send it to the renowned UK journal *Nature*. "It was a strange idea," Saar told me, "because we knew that it was practically impossible." Not only would it be their first-ever paper in such a prestigious English-language journal—in Tartu, it had long been standard to publish results in the local astronomical bulletin, the *Tartu Astrofüüsika Observatoorium Teated* (Notices of the Tartu

Astrophysical Observatory)—but at the time, the KGB thoroughly examined every piece of outgoing international mail.

The process was tedious: Simple scientific words such as "atom" needed to be avoided because secrecy-obsessed KGB censors could associate them with nuclear weapons. One of the group's corrections to the proofs was never added because their return letter to the UK was held up by the censorship process and arrived after *Nature* went to press. Nevertheless, the publication was successful. It came out in July 1974, a few months before that of the Princeton group. "Evidence is presented," they wrote, "that galaxies are surrounded by massive coronas exceeding the masses of known stars by one order of magnitude."[15] Their evidence also showed that the corona hypothesis would mean that the total mass of the universe was larger than previously thought by a factor of 10, which implied that the unknown dark matter made up the majority of matter in the universe.

From heresy to orthodoxy

On either side of the globe, Einasto's and Ostriker's groups independently demonstrated the existence of dark matter. Despite working in vastly different political contexts, both groups involved collaborations between young astrophysicists and cosmologists studying galaxies. The evidence they presented was neither a simple proof nor a single observation, like that of Zwicky or Rubin, but an inference using a combination of different arguments. As Peebles stated when I interviewed him, "What was the best argument? None of them. This is a case of no one argument being compelling, but so many arguments pointing in the same direction."

The two papers were exemplars of the nascent field of physical cosmology and its interdisciplinary teamwork and methodology: combining data and arguments from different scales—from stars and galaxies to clusters—to form a consistent physical picture of the cosmos.

YAKOV ZELDOVICH AND JEREMIAH OSTRIKER (left to right) in Moscow in 1979. (Courtesy of the AIP Emilio Segrè Visual Archives, gift of Jeremiah Ostriker.)

Despite their historical significance, the papers were not immediately received with open arms. "People thought it was just crazy," Ostriker told me. Saar recalled that "most astronomers and physicists didn't like this thing at all." Some astronomers disputed parts of the data that the groups used. Astrophysicist Geoffrey Burbidge was vocal about his disgust with the idea and authored a scathing response in the *Astrophysical Journal* a few months after the publication of the two papers. "Contrary to the results obtained by Einasto *et al.* and Ostriker *et al.*," Burbidge's paper reads, "we show that there is no unambiguous dynamical evidence which demonstrates that galaxies have very massive halos."[16] He was particularly critical of the assumption that one could measure the mass of galaxies simply by observing their dynamics.

It was only later in the 1970s that the hypothesis of missing matter—the idea that galaxies are surrounded by coronae or halos of invisible mass—became a staple in astronomical and cosmological thinking. Both the Princeton and Tartu collaborations worked hard to gain acceptance for their proposal. In 1975 Einasto organized a conference in Tallinn, Estonia, to discuss the possible

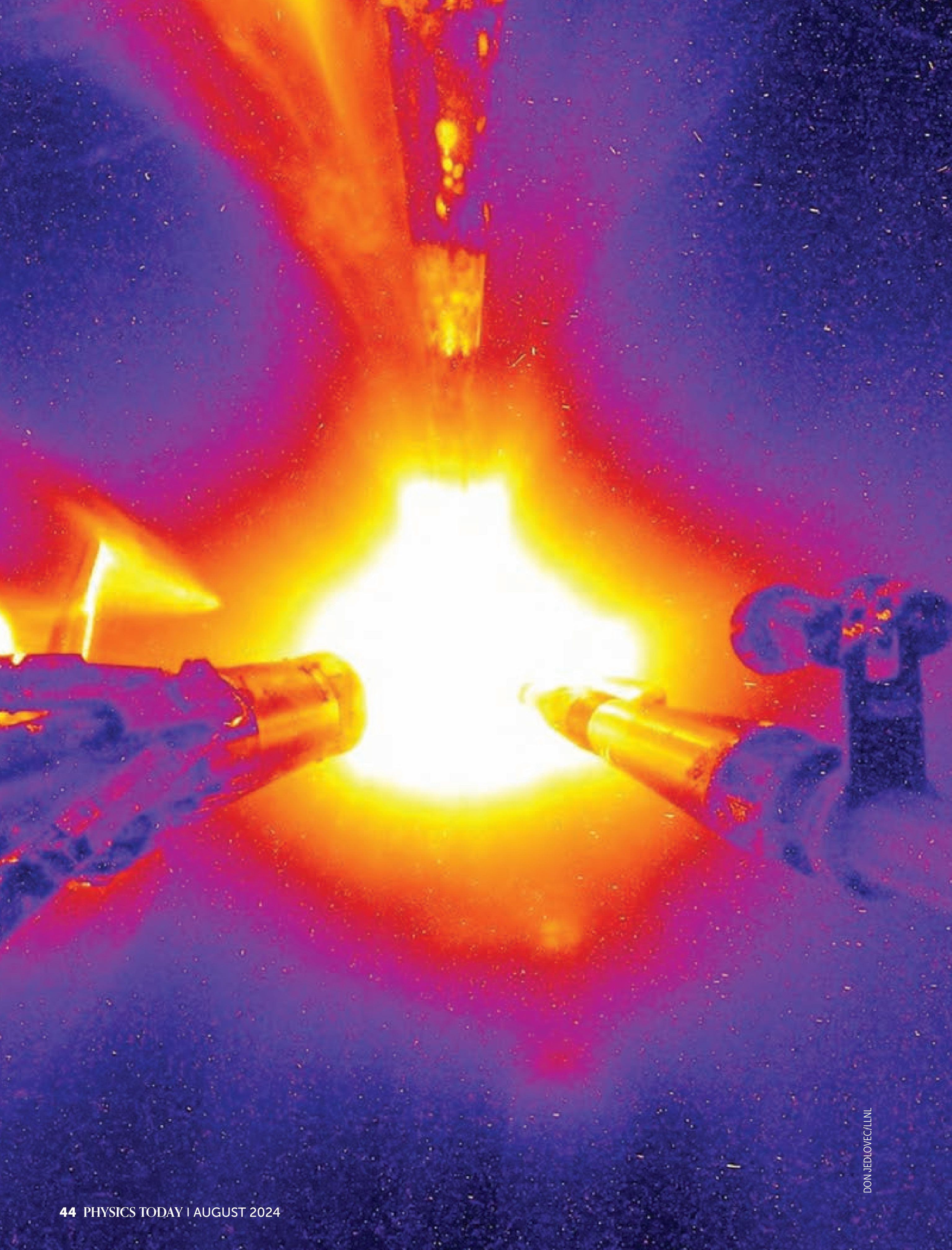
JAMES PEEBLES (far left) and Jaan Einasto (far right) converse with George Abell (second from left) and Malcolm Longair (second from right) at an International Astronomical Union symposium in 1977. (Courtesy of Jaan Einasto.)

nature of the invisible corona with Zeldovich and his students, and that same year he set up a dedicated session on "missing mass" at the Third European Astronomical Meeting in Tbilisi, Georgia. Ostriker defended his ideas in a lecture at the National Academy of Sciences in 1976, arguing that "most of the mass is not in ordinary (solar) type stars, but some other dark form."[17] By then the reception from most of the community had flipped 180 degrees: "Within two years, we went from heresy to orthodoxy," Yahil told me.

More support for their arguments appeared after 1977. Optical and radio astronomers published new galaxy-rotation data that showed more signs of unseen mass. Cosmologists began theorizing that missing matter affected galaxy formation, and particle physicists connected the mysterious substance to a potential background of neutrinos in the universe. In both instances, theorists accepted the evidence for missing mass and used the idea as a central thesis to underpin theories of cosmic particles and structure formation. In other words, what researchers soon began to call dark matter was now the basis on which theories of the universe were constructed. By the end of the 1970s, its reality appeared inescapable. Astronomers Sandra Faber and John Gallagher wrote in a 1979 review paper, "We think it likely that the discovery of invisible matter will endure as one of the major conclusions of modern astronomy."[18] Indeed it has.

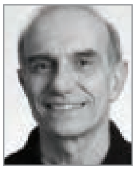
REFERENCES

1. S. Tremaine, *Astrophys. J.* **525** (centennial issue), 1223 (1999).
2. J. P. Ostriker, interview by C. Smeenk (14 May 2002), Niels Bohr Library & Archives, American Institute of Physics, College Park, MD.
3. J. de Swart, in *The Renaissance of General Relativity in Context*, A. S. Blum, R. Lalli, J. Renn, eds., Birkhäuser (2020), p. 257.
4. Unless stated otherwise, all quotations in this article come from interviews I conducted between 2014 and 2020 with Jeremiah Ostriker, James Peebles, Amos Yahil, Jaan Einasto, Enn Saar, and Alar Toomre. The transcripts will eventually be made publicly available at the American Institute of Physics's Center for History of Physics.
5. P. J. E. Peebles, *Cosmology's Century: An Inside History of Our Modern Understanding of the Universe*, Princeton U. Press (2020).
6. J. Peebles, interview by M. Harwit (27 September 1984), Niels Bohr Library & Archives, American Institute of Physics, College Park, MD.
7. J. P. Ostriker, P. J. E. Peebles, *Astrophys. J.* **186**, 467 (1973).
8. A. Toomre, in *Highlights of Astronomy, Volume 3: As Presented at the XVth General Assembly and the Extraordinary General Assembly of the I.A.U. 1973*, G. Contopoulos, ed., D. Reidel (1974), p. 457.
9. J. P. Ostriker, P. J. E. Peebles, A. Yahil, *Astrophys. J.* **193**, L1 (1974).
10. R. A. Sunyaev, in *Zeldovich: Reminiscences*, R. A. Sunyaev, ed., Chapman & Hall (2004), p. 233.
11. J. Einasto, *Dark Matter and Cosmic Web Story*, World Scientific (2014).
12. Ref. 11, p. 81.
13. J. Einasto, in *Stars and the Milky Way System*, L. N. Mavridis, ed., Springer (1974), p. 291.
14. Ref. 11, p. 90.
15. J. Einasto, A. Kaasik, E. Saar, *Nature*, **250**, 309 (1974); "Dynamical evidence of massive coronas of galaxies," *Tartu Astrofüüs. Obs. Publ.* preprint (1974).
16. G. Burbidge, *Astrophys. J.* **196**, L7 (1975).
17. Ref. 11, p. 93; E. K. Kharadze, ed., *Proceedings of the Third European Astronomy Meeting, Tbilisi, 1–5 July 1975: Stars and Galaxies from Observational Points of View*, Academy of Sciences of the Georgian SSR (1976); J. P. Ostriker, *Proc. Natl. Acad. Sci. USA* **74**, 1767 (1977), p. 1773.
18. S. M. Faber, J. S. Gallagher, *Annu. Rev. Astron. Astrophys.* **17**, 135 (1979), p. 182.



DON JEDLOVEC/LLNL

Stefano Atzeni is a consulting senior scientist at Focused Energy GmbH. He retired in 2023 as a physics professor at Sapienza University of Rome. He is a coauthor of *The Physics of Inertial Fusion* (2004). **Debra Callahan** was one of the scientific leaders for the design that achieved ignition at the National Ignition Facility. She recently joined Focused Energy to use lessons from NIF results to develop a fusion power plant.



HARNESSING ENERGY FROM LASER FUSION

Stefano Atzeni and Debra Callahan

A new goal for nuclear fusion is underway: Commercialize the nascent technology that first demonstrated net energy gain in 2022 at the National Ignition Facility.

Researchers in nuclear fusion reached a critical milestone on 5 December 2022. On that day, Lawrence Livermore National Laboratory's National Ignition Facility (NIF) demonstrated for the first time that a controlled nuclear fusion reaction could produce more energy than the energy put into it. The reaction yielded 3.1 MJ of energy from an energy input of 2 MJ.[1] In 2023, a similar experiment released 3.9 MJ of energy for the same energy input.

The accomplishments at NIF were made possible by inertial confinement fusion (ICF).[1] In ICF a powerful energy source, typically a laser, is used to heat and compress fuel to such a high temperature and density that the fuel cannot be equilibrated by any mechanical or electromagnetic forces. Therefore, once compressed, the fuel remains confined only because of its own inertia, and that's when the fusion reaction takes place. The fuel stays compressed until it's naturally uncompressed.

The time that takes is defined by the movement of a rarefaction wave across the fuel target, which typically takes less than a nanosecond.[2,3] The entire ICF sequence is pulsed, and for energy production, the sequence has to be repeated cyclically. Each target must release an amount of energy much more than that delivered to it. For the NIF achievement, the reacting fuel, a mixture of deuterium–tritium (DT) in a plasma state, was heated to a high temperature and confined at a high density for a sufficiently long time so that the energy it produced exceeded the energy delivered by the laser.[4]

Fusion researchers have also been studying magnetic confinement fusion, which uses magnetic fields that are 100 000 times as strong as Earth's. The strong fields keep hot plasmas in equilibrium inside a vacuum chamber.[4] Large national and international programs have supported research in magnetic confinement fusion since the late 1950s and operate experimental devices with different magnetic field configurations. The best-performing ones so far are torus-shaped devices known as tokamaks. See the box on page 49 for more details on magnetic confinement fusion.

The net-gain milestone from NIF has ignited a push to generate commercial-scale fusion energy with the ICF approach.

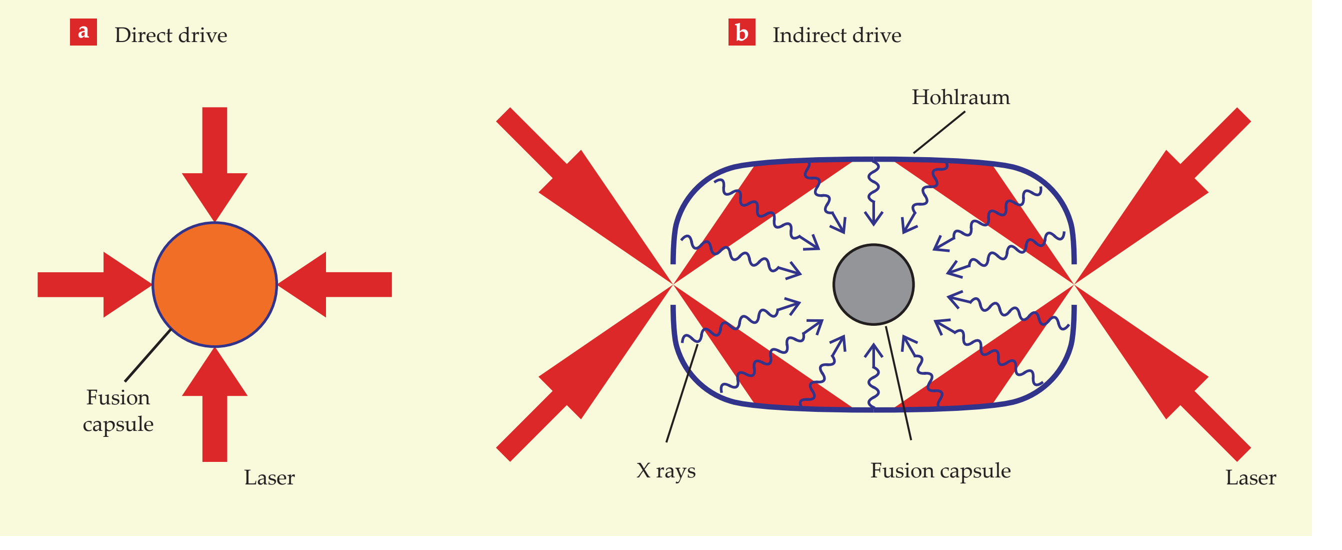


FIGURE 1. INERTIAL CONFINEMENT. (a) In a direct-drive approach, a nanosecond pulse of laser light drives a capsule of deuterium–tritium fuel to heat up and compress. The fuel's own inertia keeps it confined long enough to become a burning plasma and sustain a fusion reaction. **(b)** An indirect-drive approach works similarly but involves applying laser-generated thermal x rays in a cylindrical cavity, known as a hohlraum, to deliver the energy to the fusion capsule. (Adapted from S. Atzeni, *Plasma Phys. Control. Fusion* **51**, 124029, 2009.)

Now that some of the fundamental physics issues have been overcome, public and private enterprises are working on solving the technological challenges faced to use inertial fusion to produce electricity.

INGREDIENTS AND ISSUES

For fusion reactions inside stars, confinement is caused by gravitational forces. In 1920, before nuclear reactions and neutrons were discovered, Arthur Eddington hypothesized that the energy source of stars could be the fusion of hydrogen atoms to form a helium atom. By the early 1930s, fusion reactions were detected in laboratory experiments, and a few years later, they were acknowledged as the source of stellar power.

The Sun, for example, is powered by a series of fusion reactions. It starts with the fusion of two protons into a deuteron—a type of beta decay based on the weak interaction—and ends with the formation of helium. So although the series of reactions is different from Eddington's hypothesis, his determination of the initial and final products is essentially correct.

The oft-repeated motto of "harnessing the power of the stars" is evocative but reductive. Terrestrial fusion-energy production demands that the hydrogen with just one proton be replaced with a fuel mixture of its isotopes deuterium and tritium, which is 10^{25} times as reactive. In a fusion reactor, the DT reaction produces an alpha particle and a neutron. Deuterium can be extracted inexpensively from ordinary water, but tritium exists only in trace amounts. Neutrons produced in the fusion reaction, when reacted with lithium, can provide the necessary tritium for the DT reaction. To produce power, the DT fuel needs to reach a temperature of 50–100 million K, which is several times as high as the temperature at the center of the Sun.

The extremely large fuel-specific yield of DT (334 TJ/kg) limits the fuel's mass in a target to a few milligrams. Otherwise, the explosive energy could not be contained in the reactor. A large gain from such a small amount of fuel can be obtained only if the fuel is compressed to a density about 1000 times as high as that of liquid DT. To achieve a high gain $G \gg 1$—with G being the ratio of fusion energy to energy delivered to a fusion target—a laser must heat a small region in the fuel until it ignites. Once the hot spot is generated and DT reactions occur, most of the neutrons from the fusion reactions escape unscattered. But the high-density DT fuel slows the alpha particles, which then deposit their energy into the fuel.

When the power that the alpha particles deposit in the hot spot exceeds that lost by it—because of bremsstrahlung radiation, electron thermal conduction, and mechanical work—the hot spot self-heats, and a burn wave propagates through the fuel. For the hot spot to ignite, several conditions must be met. Its temperature must be above 5 keV (about 50–60 million K), and a substantial part of the energy of the alpha particles must be contained in the hot spot. Additionally, the hot spot needs to have a pressure twice that at the center of the Sun.

Such extreme thermodynamic conditions can be obtained with laser-driven implosion, a scheme devised immediately after the invention of the laser and first discussed in the scientific literature a decade later.[5] Figure 1 shows how a nanosecond pulse of either laser light (direct drive) or laser-generated thermal x rays (indirect drive) delivers energy uniformly to the surface of a millimeter-sized hollow spherical capsule.[6] The capsule consists of an outer layer of radiation-absorbing material, known as the ablator, and an inner layer of frozen DT fuel.

In both direct-drive and indirect-drive approaches, the outer portion of the capsule is heated and ablated and forms a hot expanding plasma, as shown in figure 2. Because of the 100 Mbar pressure exerted by the expanding plasma, momentum conservation induces the remaining portion of the shell to implode like a spherical rocket at 300–500 km/s. About 5–15% of the energy absorbed by the capsule is transmitted to the fuel in the form of kinetic energy. As the implosion stagnates, kinetic energy is converted into internal energy, the fuel is compressed, and a hot spot forms at its center.

The spherical-rocket mechanism concentrates energy in space and amplifies the pressure generated by the ablating plasma from hundreds of megabars on the capsule's surface to hundreds of gigabars in the compressed fuel when the capsule stagnates. Its realization, however, requires researchers to address several critical issues concerning energy efficiency and implosion symmetry and stability.[7] The efficient coupling of laser energy to the fuel requires short-wavelength light in the visible or UV range and with a peak intensity that is large enough to generate the necessary 100 Mbar pressure but not high enough to excite deleterious plasma instabilities that degrade light absorption.

In addition, the power of the laser pulse must be shaped in time to ensure a smooth, low-entropy compression of most of the fuel. Spherical convergence of the fuel and the generation of a central hot spot demand a highly symmetrical implosion and an almost perfectly uniform spherical irradiation. Hydrodynamic instabilities, which are seeded by unavoidable imperfections in the target and irradiation nonuniformities at short spatial and temporal scales, must be considered too. The instabilities can threaten the integrity of the capsule's shell and can cause undesired materials to mix with the fuel. The thickness of the shell, or more precisely, its ratio of radius to thickness, is a crucial parameter. Its value, as with many other design choices, is a compromise: Although a thicker target is less sensitive to asymmetries and instabilities, it requires more intense irradiation, making it more susceptible to plasma instabilities.

DEMONSTRATING IGNITION

In the early 1990s, scientists at Lawrence Livermore National Laboratory were confident that ignition could be demonstrated. They had accumulated considerable theoretical, experimental, and computational data over two decades. The national-lab scientists chose indirect drive as the main approach instead of direct drive,[6] because they believed the former allowed better control of the fuel capsule's symmetry and stability.

The laser pulse and target were designed to control the crucial issues of symmetry and stability. The target, a 1-cm-long cylindrical cavity known as a hohlraum, encloses a hollow spherical fusion capsule, as shown in figure 3. The target is irradiated by the 192 beams created at NIF. Together, they deliver 15–25 nanosecond pulses of 350 nm UV light with a total energy of 2 MJ and a peak power of 500 TW. Inside the hohlraum, the laser light is converted to x rays, which then indirectly drive the capsule implosion.

NIF experiments began in 2009, and a first ignition campaign was performed in the following three years. The results were disappointing: Fusion yields measured a few kilojoules instead of the expected megajoules. An analysis showed that a

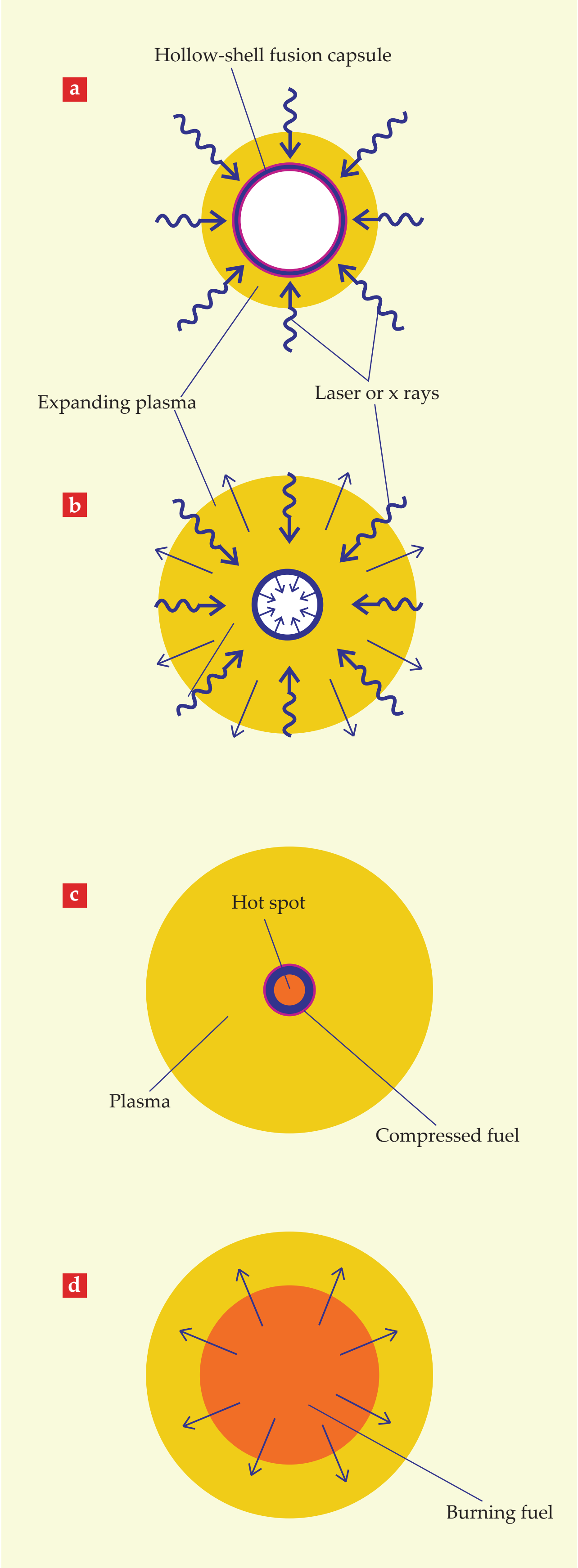


FIGURE 2. LASER FUSION. After a spherical capsule of fusion fuel is **(a)** irradiated by a laser system, the outer layers of the shell are heated and form a hot expanding plasma (yellow). **(b)** The huge pressure from the plasma forces the remaining part of the capsule, with the deuterium–tritium fuel, to implode. **(c)** The result is the formation of a hot spot, which is surrounded by denser, colder fuel. **(d)** Once ignited, the burning fuel expands outward, and the released fusion energy is harnessed to generate electricity. (Adapted from S. Atzeni, *Plasma Phys. Control. Fusion* **51**, 124029, 2009.)

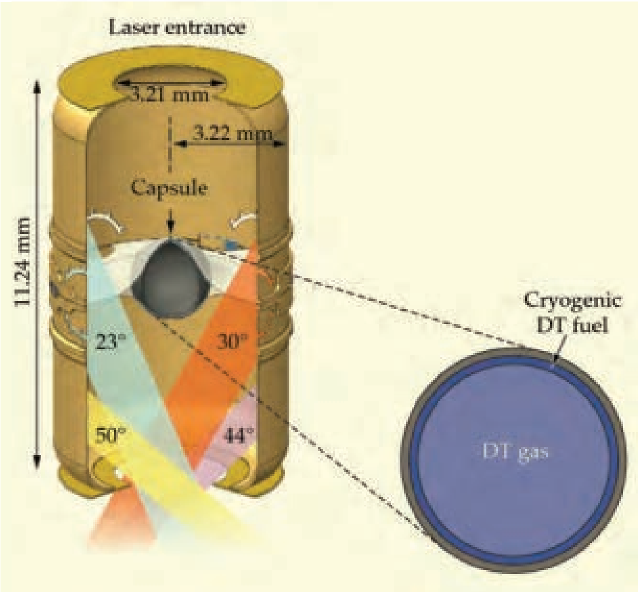


FIGURE. 3. THE FUSION TARGET that was used to demonstrate ignition in National Ignition Facility experiments consists of a high-density carbon capsule filled with deuterium–tritium gas with a cryogenically cooled outer shell of DT fuel. The capsule is suspended in a cylindrical cavity known as a hohlraum. It converts the laser light to x rays, which drive the capsule implosion. (Adapted from ref. 4.)

few discrepancies in the simulations used to design the experiments were to blame for the poor performance. For instance, the plasma generated inside the hohlraum caused reduced laser-coupling efficiency, deleterious laser–plasma interactions,[8] and a less symmetric implosion.

The hydrodynamic instabilities seeded by small engineering features in the design were another problem. The 10-μm-wide tube used to fill the capsule with DT fuel and the thin membrane that keeps the capsule at the center of the hohlraum, for example, caused carbon atoms from the ablator to mix with the DT fuel. The mixing enhanced radiation losses and hindered hot-spot heating.

In the following years, the fusion program at NIF—carried out by roughly 1000 scientists from the US and around the world—made a series of changes. They refined the original design to progressively control the implosion without altering the general scheme.[9] The team's many significant changes included altering the laser pulse's shape, modifying the dimensions of the hohlraum and capsule, replacing the hohlraum material and the ablator material, and reducing the size of the fuel-filling tube and the membrane that holds the capsule. In addition, the laser pulse energy was raised from 1.9 to 2.1 MJ, and power distribution between the beams was improved through the exploitation of nonlinear laser–plasma interactions.

The efforts dramatically improved the setup and culminated in several achievements: in 2021, hot-spot ignition (see figure 4) and the release of 1.3 MJ of energy;[10] in 2022, a first demonstration of gain that exceeded unity with $G = 1.5$ and a yield of 3.1 MJ; and in 2023, the release of 3.9 MJ.[1] (For more on the first experiment to achieve fusion energy gain, see "National Ignition Facility surpasses long-awaited fusion milestone," Physics Today online, 13 December 2022.)

The demonstration of ignition at NIF proves the soundness of both laser-driven compression and hot-spot ignition followed by burn propagation—two of the basic pillars of any laser inertial fusion scheme. The demonstration changes the prospects of inertial fusion research and opens a path toward inertial fusion energy (IFE). Although energy production was always considered a possible long-term goal of ICF research, the US program and NIF were funded by the Department of Energy's National Nuclear Security Administration as part of the US Stockpile Stewardship Program. The achievement of ignition, however, has triggered various new initiatives, including an increase in the number of private companies that are pursuing IFE.

INERTIAL FUSION ENERGY

Major progress in many areas is still needed to make IFE a reality. In a laser-IFE reactor, the target's release of fusion energy is eventually transformed into electrical energy, typically with a steam turbine system, with a thermal efficiency of 40%. A fraction f of the energy is used to energize the laser, while the complementary fraction $1 - f$ is delivered to the grid.

Preliminary system studies and economic analyses indicate that a value of $f = 0.25$ is the maximum amount of energy that can be used to energize the laser while still being compatible with economically viable energy production. To reach that goal, the product of target gain and laser efficiency must be larger than about 10. For a 10% efficient laser, that corresponds to a target gain of about 100, which is 50 times as large as what has been demonstrated at NIF. So far, NIF has ignited a few targets in single-shot experiments. But a commercial-sized gigawatt electric power plant would need to operate at a high repetition rate, with a few low-cost targets ignited per second.

Are such tremendous improvements possible, and if so, on what time scale? Before we consider those questions, remember that the NIF laser is based on 30-year-old technology, and the laser architecture and targets were designed to maximize the probability of ignition with the existing technology. The progress in fusion research over the past few decades reveals many options for improvement.

On the target side, NIF has so far demonstrated $G = 1.9$, which was about a factor of 30 increase compared with the G value achieved at the end of 2020. An additional increase, by a factor of 5, may be possible with the current indirect-drive scheme at NIF. An improvement by a factor of 5–10 in the coupling efficiency is possible by the replacement of indirect drive with direct drive. Progress in laser and target technology has overcome some of the limitations that led to the decision by NIF scientists to opt for an indirect-drive architecture.

Modern lasers can produce the very smooth, incoherent beams that are required to reduce the effect of both laser-induced perturbations of the surface target, which seed hydrodynamic instabilities, and laser-plasma instabilities, which reduce the absorption of laser light. Experiments at the Omega laser at the Laboratory for Laser Energetics in Rochester, New York, have recently shown that a central hot spot can be efficiently formed,[11] and models indicate that ignition could be achieved with the use of larger targets and lasers.

Another improvement by a factor of 2–3 can be attained by using an advanced ignition scheme, such as shock ignition[12] or fast ignition.[13,14] Such schemes still involve laser-driven implosion, but the stages of compression and hot-spot ignition are partly separated. A fusion target with a thick layer of DT fuel is first compressed to high density at as cold a temperature as possible and at a lower implosion velocity than in the conventional approach. If the velocity is lowered by half, then four times the amount of fuel can be used for the same kinetic energy that compresses the fuel. The lower velocity and thicker fuel reduce the risk of instabilities and make a more efficient use of the laser energy.

Larger targets with thicker fuel layers take longer to explode, which increases the burn-up fraction from about 5%, which is what's been demonstrated at NIF, to 20–30%. The reduced implosion velocity, however, prevents the formation of an igniting hot spot. In fast ignition, a hot spot is generated in the compressed fuel by electrons or ions that are accelerated by an additional ultraintense multipetawatt laser pulse that lasts for a few picoseconds (see figure 5). In shock ignition, the hot spot forms from the nearly spherically symmetric irradiation of the shell by a multibeam laser pulse with a power of up to 1 PW and a duration of a few hundred picoseconds.

On the laser side, NIF has used a low efficiency flashlamp-pumped glass laser, which is limited to single-shot operation. But it can be replaced by a high-efficiency diode-pumped glass laser. The technology—demonstrated so far in small devices—needs to be scaled to megajoule systems. The rapidly decreasing costs of diodes make such upscaling conceivable to us in the next 10–15 years. Excimer lasers are an alternative to solid-state lasers. They use a gas as the amplifier medium and have the advantage of not damaging the amplifier material.[15] In addition, because of their shorter wavelengths, excimer lasers generate higher pressures than solid-state lasers with the same intensity.

Fuel targets are currently manufactured by hand for experiments because only a small number of targets of any one design are being produced. Each one is carefully characterized in a labor-intensive process. But for a commercial power plant, up to a million targets have to be produced, characterized, and injected into the reactor vessel each day. That level of scaling will only be possible with commercial-scale mass production. The approach could significantly reduce the cost to make targets, but it has yet to be demonstrated.

A magnetic fusion energy device consists essentially of a single complex machine, in which the plasma chamber is surrounded by heat-exchange circuits, a tritium-breeding blanket, and superconducting magnets. The IFE approach, however, has its lasers and other main components in a separate building far from the harmful reactor environment. The separation allows not only for maintenance when the reactor is off but also for a wider choice of materials for the reactor. Although an IFE reactor is much simpler than a magnetic-fusion-energy one, it has to provide analogous functions concerning heat transfer, tritium breeding, and radiation containment. Some IFE-specific issues that still need to be addressed are target insertion and tracking and the optimization of the interface between the reactor and the laser optics.

AN ACCELERATED PATH

Of the many companies that have been established over the past 10 years to commercialize fusion energy, just a few are

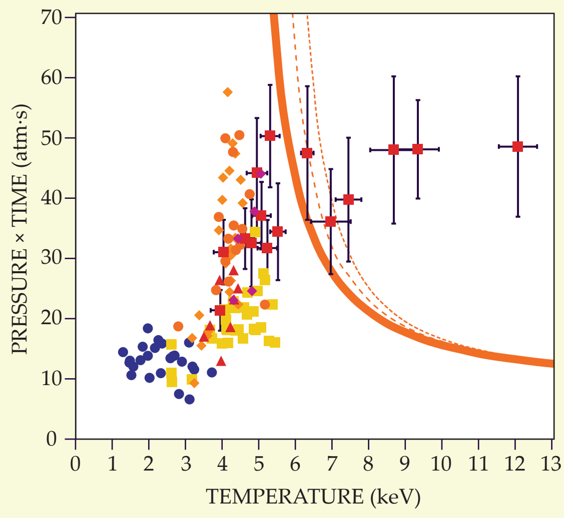


FIGURE 4. PERFORMANCE OF FUSION EXPERIMENTS. The product of a hot spot's peak pressure and measured burn time in experiments is plotted as a function of the hot spot temperature (1 keV = 1.16 × 10^7 K). Ignition has occurred if a data point and its uncertainty bars are above the solid orange curve. For experiments in which deuterium–tritium fuels are contaminated by carbon impurities, ignition has occurred if a data point is above the orange dashed curves. Three experiments overcame the ignition boundary in 2021–22, and three others overlapped with the boundary but did not demonstrably exceed it, given their uncertainties. The oldest experiments, from 2010, are plotted in the bottom-left corner, and the data show a steady improvement. (Adapted from ref. 4.)

Magnetic confinement fusion

Some major recent achievements in magnetic fusion energy include the operation of a steady-state plasma for 1000 seconds at the Experimental Advanced Superconducting Tokamak in Hefei, China,[17] and the production of 69 MJ of energy in 5.2 seconds at the Joint European Torus in Culham, UK.[18] Among magnetic-fusion machines, JET holds the record value of 0.67 for *Q*, the ratio of fusion thermal power to the power injected into the machine to heat the plasma. The injected power takes the form of electromagnetic waves and energetic particle beams.

Higher values of *Q* are necessary for commercial-scale fusion power. The goal of ITER—a consortium of countries constructing a large tokamak in France—is to produce 500 MW of fusion power in runs of a few minutes with *Q* values of 5–10. The ITER tokamak is scheduled to start operating in 2034, and DT fuel may be used a few years later. Many ITER partners have begun conceptual designs and technological development of successor reactors, known as DEMOs. A few private companies have raised some $4 billion, according to the Fusion Industry Association, to pursue accelerated paths to fusion energy that use magnetic confinement.

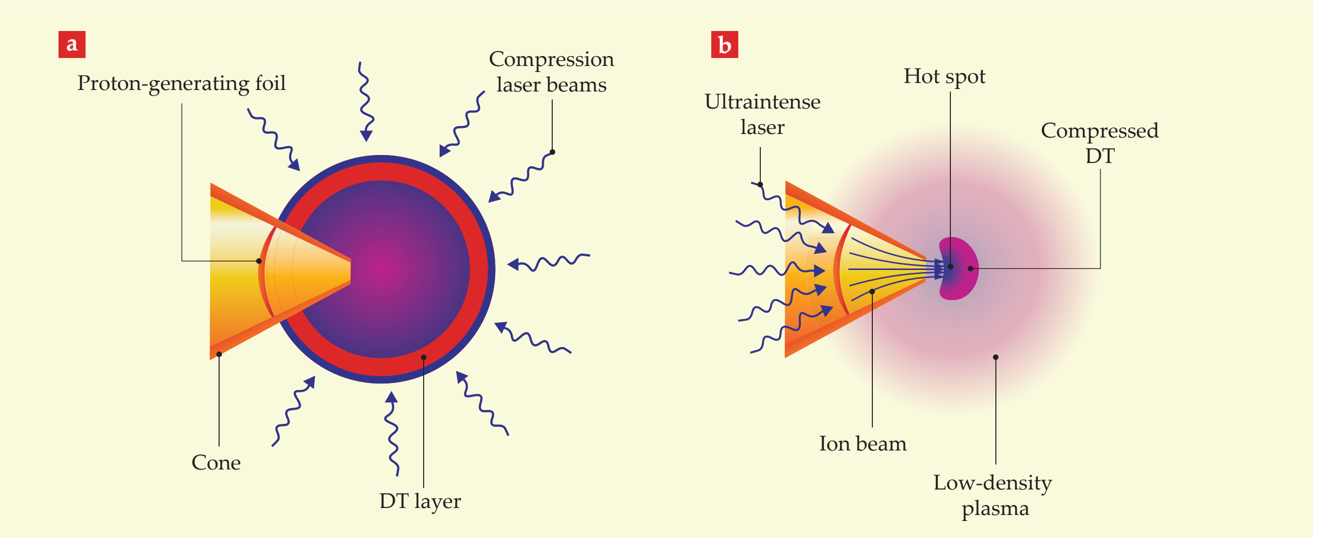


FIGURE 5. PROTON FAST IGNITION. (a) When spherical laser irradiation hits a fusion-fuel target, the indirect the outer shell is heated and blown away, which triggers the target to implode, at a velocity of about 200 km/s. **(b)** A proton beam is generated by the interaction of an ultraintense laser beam with a proton-generating foil supported by a cone. The ion beam then heats the compressed fuel to create a hot spot. The cone provides a clean path to the laser beam and shields the proton-generating foil from the expanding plasma and its radiation. In such a hot, compressed state, the hydrogen isotopes deuterium and tritium can fuse into helium, and the fusion energy can be harnessed as electrical energy. (Courtesy of Kateryna Bielka, Focused Energy.)

focusing on laser-driven IFE (see "The commercial drive for laser fusion power," Physics Today online, 20 October 2021). Each company is pursuing a different approach to achieve the high gain needed for IFE. A feature that all the approaches share is modularity. Individual components, such as the laser or the target, can be developed and tested in parallel and even exchanged in a combined system when better solutions are available.

Among US IFE startups, Xcimer Energy uses an argon fluoride excimer laser and a hybrid of indirect and direct drive, and Pacific Fusion is pursuing a magnetic-drive approach. The German–US company Focused Energy—which we both work for—pursues laser direct drive and uses diode-pumped solid-state lasers with proton fast ignition, although shock ignition is being evaluated as a backup solution.

Focused Energy is developing a modular, stepwise approach. A scaled-down facility with a limited number of beams could be built first, and then laser systems could be added over time to increase performance. The company plans to start with a small laser system within the next two years to address important physics questions. Then it will develop a subscale implosion facility where, without igniting a target, all the technology components can be developed to the required readiness level. They'll be assembled in a demonstration fusion power plant that is slated to begin operations in about 15 years. The driving laser systems, the target technology, and the diagnostics systems can thereby be tested before the first run, which will save money and time.

Public institutions have been receptive to the entry of private IFE companies into the field. National labs, private companies, and universities have developed private–public partnerships in the US,[16] Germany, and the UK. In addition, the Department of Energy's LaserNetUS program provides academic and private users access to laser facilities. Private companies can benefit from the body of expertise accumulated by public institutions over decades of fundamental research. Public institutions, in turn, could learn about more agile initiatives and high-risk, high-reward innovative concepts.

The results from NIF demonstrate that a laser-driven fusion target can achieve a target gain $G > 1$ and that the basic physics is understood. Although challenges remain for researchers and companies to take the NIF results and produce an economical power plant, the path to fusion energy is clear.

REFERENCES

1. H. Abu-Shawareb et al. (Indirect Drive ICF Collaboration), *Phys. Rev. Lett.* **132**, 065102 (2024).
2. S. Atzeni, J. Meyer-ter-Vehn, *The Physics of Inertial Fusion: Beam Plasma Interaction, Hydrodynamics, Hot Dense Matter*, Oxford U. Press (2004).
3. R. Betti, O. A. Hurricane, *Nat. Phys.* **12**, 435 (2016).
4. J. P. Freidberg, *Plasma Physics and Fusion Energy*, Cambridge U. Press (2007).
5. J. Nuckolls et al., *Nature* **239**, 139 (1972).
6. J. D. Lindl, *Inertial Confinement Fusion: The Quest for Ignition and Energy Gain Using Indirect Drive,* AIP Press (1998).
7. S. E. Bodner, *J. Fusion Energy* **1**, 221 (1981).
8. P. Michel, *Introduction to Laser–Plasma Interactions*, Springer (2023).
9. A. L. Kritcher et al., *Phys. Rev. E* **109**, 025204 (2024).
10. H. Abu-Shawareb et al. (Indirect Drive ICF Collaboration), *Phys. Rev. Lett.* **129**, 075001 (2022).
11. C. A. Williams et al., *Nat. Phys.* **20**, 758 (2024).
12. R. Betti et al., *Phys. Rev. Lett.* **98**, 155001 (2007).
13. M. Tabak et al., *Phys. Plasmas* **1**, 1626 (1994).
14. M. Roth et al., *Phys. Rev. Lett.* **86**, 436 (2001); S. Atzeni, M. Temporal, J. J. Honrubia, *Nucl. Fusion* **42**, L1 (2002).
15. S. P. Obenschain et al., *Philos. Trans. R. Soc. A.* **378**, 20200031 (2020).
16. S. C. Hsu, *J. Fusion Energy* **42**, 12 (2023).
17. Y. Song et al., *Sci. Adv.* **9**, eabq5273 (2023).
18. "Breaking new ground: JET tokamak's latest fusion energy record shows mastery of fusion processes," *EUROfusion News*, 8 February 2024.

PT

NEW PRODUCTS

Focus on software, data acquisition, and instrumentation

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

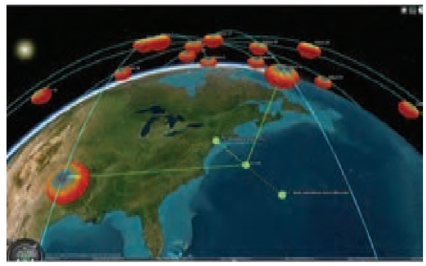
Andreas Mandelis

Low-pulsation pumps

KNF has expanded its Smooth Flow pump series by adding two customizable models, the FP 7 and FP 25, which are suitable for demanding applications. With a built-in dampener to provide smooth, low-pulsation flow that minimizes friction in tubing and prevents bubble formation, the FP 7 and FP 25 reliably transfer fluids without the need for additional equipment. They feature four-point valves for reliable self-priming, even at reduced motor speeds, and a digitally adjustable brushless DC motor for precise pump control and pinpoint matching to individual electronics. The FP 7 has a nominal flow rate of 70 ml/min, a minimum suction height of 2 m of water column (mH_2O), and a maximum pressure of 10 mH_2O. The FP 25 offers a nominal flow of 250 ml/min and a minimum suction height of 3 mH_2O. The pumps are also available in high-pressure versions: both the FP 1.7 and FP 1.25 achieve nominal pressures up to 60 mH_2O. ***KNF Neuberger Inc***, *2 Black Forest Rd, Trenton, NJ 08691, https://knf.com*

Signal generation software

Siglent designed its SigIQPro software to simplify general and standards-based signals creation for the development and verification of wireless signal devices. The PC-based software generates a rich variety of IQ (quadrature) signals and supports Bluetooth, internet of things, custom orthogonal frequency-division multiplexing, and more. It simplifies the test setup, reduces simulation time, and lowers test costs. Users need only set the basic parameter information of the signals to generate the required waveforms quickly and easily. SigIQPro's user-friendly interface provides insights and displays a wide array of performance metrics. The software is compatible with the company's SSG5000X-V series vector signal generator and SDG7000A series multifunction dual-channel arbitrary waveform generator. Waveforms can be transferred directly to the signal generators via GPIB, USB, and LAN. ***Siglent Technologies NA Inc***, *6557 Cochran Rd, Solon, OH 44139, https://siglentna.com*



Mathematical programming software

New functionality in Release 2024a (R2024a), an update to MathWorks' MATLAB and Simulink software, streamlines the development of artificial intelligence and wireless communications systems. The Satellite Communications Toolbox, which provides standards-based tools for designing, simulating, and verifying satellite communications systems and links, has been updated to enable users to model scenarios. It can be used to design physical-layer algorithms for RF components and ground station receivers, generate test waveforms, and perform golden reference design verification. R2024a's major updates to MATLAB and Simulink tools include the Computer Vision Toolbox, which provides algorithms, functions, and apps for designing and testing computer vision, 3D vision, and video processing systems; the Deep Learning Toolbox for designing and implementing deep neural networks; and the Instrument Control Toolbox, which connects MATLAB directly to instruments such as oscilloscopes, function generators, signal analyzers, power supplies, and analytical instruments. ***The MathWorks Inc***, *1 Apple Hill Dr, Natick, MA 01760, www.mathworks.com*

Compact parallel-kinematic multiaxis motion platform

Physik Instrumente (PI) has unveiled the A-361, a planar XY-Theta-Z nanopositioning stage based on air bearings and direct drive motors. Unlike conventional multiaxis stages, the A-361 uses only one moving platform shared among its three axes. Despite featuring an integrated vacuum chuck to accommodate wafers or other samples, the stage has an ultralow profile of only 40 mm. With its noncontact components, the A-361 ensures frictionless operation and thereby offers ultrahigh precision, superior performance, and a maintenance-free, virtually infinite service life. The built-in position-locking capability is useful for operations that require high stability with zero servo jitter. The contact-free operation and lack of lubricants make the stage suitable for clean-room applications. If precision motion in six axes is needed, as is often the case in photonics alignment and semiconductor inspection applications, the A-361 stage can be integrated with the company's A-523 Z Tip/Tilt stage to create a six-degrees-of-freedom nanopositioning system only 100 mm high. ***Physik Instrumente LP***, *16 Albert St, Auburn, MA 01501, www.pi-usa.us*

Python package for digitizers and generators

Spectrum Instrumentation developed its spcm open-source Python package to make programming its current test and measurement instruments faster and easier. Python is simple and versatile and offers an extensive collection of libraries and frameworks, such as NumPy, that can accelerate programming development cycles. The spcm package lets users take advantage of the Python language by providing a high-level, object-oriented programming interface designed for the company's digitizer, arbitrary waveform generator, and digital input–output products. It includes the full source code and detailed examples. The Python package safely handles the automatic opening and closing of cards, groups of cards, and Ethernet instruments and the allocation of memory for transferring data to and from those devices. It offers sampling rates from 5 MS/s to 10 GS/s and supports the use of real-world physical quantities and units, such as 10 MHz, so users can directly program driver settings in their preferred unit system. Available on GitHub, spcm is free of charge under the MIT license. ***Spectrum Instrumentation Corp***, *401 Hackensack Ave, 4th Fl, Hackensack, NJ 07601, https://spectrum-instrumentation.com*

Quantum parameter measurements

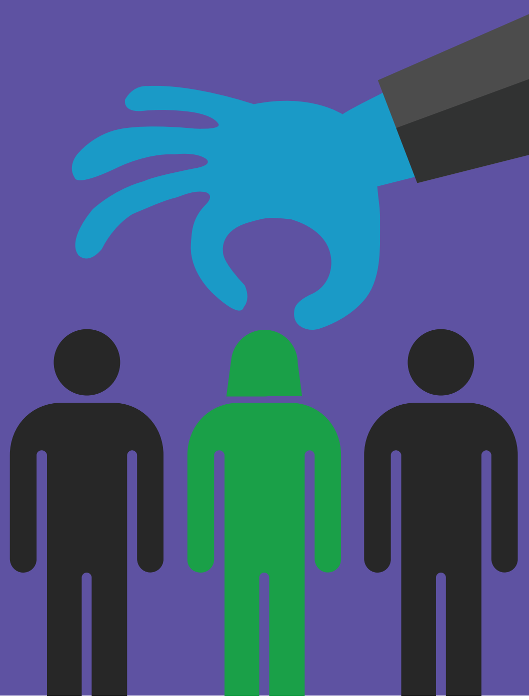
Keysight has introduced QuantumPro, which the company says is the electronic design automation industry's first integrated workflow tailored for the seamless design of superconducting qubits. QuantumPro consolidates five essential functionalities—schematic design, layout creation, electromagnetic analysis, nonlinear circuit simulation, and quantum parameter extraction—into Keysight's PathWave Advanced Design System 2024 software platform. It translates traditional microwave outputs into easily adjustable quantum parameters. The electromagnetic solution uses the method of moments to provide a cost-effective simulation that can accelerate the design cycle. Instead of solving for the electric field in the full volume, the method of moments solves only for the currents on the metal surface. Co-simulation and parametric analysis enable seamless tuning of qubits and resonators. The new release also includes an alpha feature for a Python console that lets users control workflow, modify workspaces, automate tasks, and customize the user interface. ***Keysight Technologies Inc***, *1400 Fountaingrove Pkwy, Santa Rosa, CA 95403, www.keysight.com* PT

LOOKING FOR A JOB?

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

LOOKING TO HIRE?

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



Questions? Email us at ptjobs@aip.org

PHYSICS TODAY | JOBS

QUICK STUDY

Brian Arbic is a professor of physical oceanography at the University of Michigan in Ann Arbor. **Brandon Johnson** is a professor of geophysics and planetary sciences at Purdue University in West Lafayette, Indiana. **Vasily Titov** is chief scientist at the NOAA Center for Tsunami Research in Seattle, Washington.

The tsunami triggered by the Chicxulub impact

Brian K. Arbic, Brandon C. Johnson, and Vasily Titov

The large wave caused by the dinosaur-killing asteroid impact eroded the seafloor and disturbed sediments in ocean basins far from the collision site.

An asteroid struck the ocean's continental shelf near the modern-day Mexican town of Chicxulub some 66 million years ago. The collision energy was 3×10^{23} J, which is equivalent to about 5 billion Hiroshima-sized atomic bomb explosions. Within seconds, the impact excavated a hole more than 20 km deep into Earth's surface and threw bedrock and sediment into the upper atmosphere.

As impact ejecta fell back to Earth, the heat from their reentry generated a global pulse of thermal radiation similar to the conditions of an oven set on broil for about one hour. The thermal radiant flux and hot falling ejecta ignited widespread forest fires. The resulting soot blocked sunlight for several years, and the ensuing frigid surface temperatures led to widespread plant die-offs and the demise of a vast array of animals. Approximately 75% of species, most famously the non-avian dinosaurs, went extinct shortly after impact. Other life forms, including mammals and flowering plants, reemerged as dominant species.

Scientists in the 1980s and 1990s pinpointed the crater location from anomalies in Earth's gravity and magnetic fields. More recently, measurements from drill cores corroborated the basic conclusions of crater-formation models. They also suggested that the impact could have triggered a powerful tsunami. Motivated by that possibility, the three of us and several colleagues asked whether such a tsunami could have had global effects. In 2022 we published the first global simulation of the Chicxulub megatsunami.

The first 10 minutes after the initial impact and the tsunami's formation were simulated with a shock-physics model that's widely used to study crater formation on Earth and other bodies in the solar system. The results of that model were then fed into propagation models to track the tsunami across the globe for two days. Finally, we examined geological records for evidence of global tsunami propagation.

Simulating the impact

The cavity formed by the impact was initially 28 km deep, but under the force of gravity, it collapsed into a crater roughly 3 km deep. As the crater took shape, a curtain of ejecta pushed water outward and initiated the advancing rim wave, shown in blue in figure 1. Some 10 minutes after impact, the wave had spread more than 200 km from the collision site, its width had grown to 50 km, and its height was 1.5 km. The energy of the impact tsunami was about 30 000 times as great as that of any tsunami generated by a large earthquake today, including the 2004 Boxing Day tsunami that originated off the coast of Sumatra in the Indian Ocean (see Physics Today, June 2005, page 19) and the 2011 Tohoku tsunami generated near Japan (see the article by Thorne Lay and Hiroo Kanamori, Physics Today, December 2011, page 33).

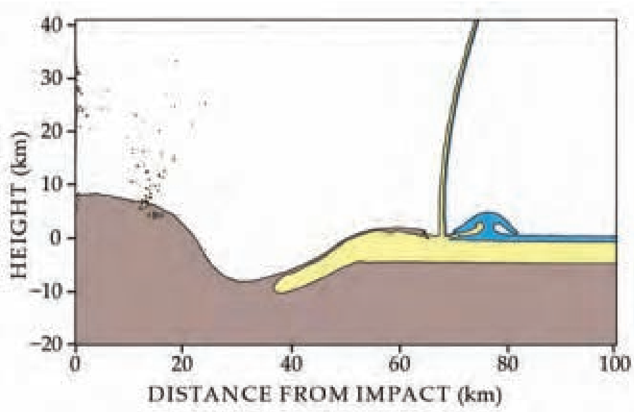


FIGURE 1. THE CHICXULUB IMPACT 66 million years ago displaced crustal material (brown) and sediments (yellow) and launched a tsunami that spread around the globe. The origin marks the point of impact. A curtain of ejected sediment, tens of kilometers high, pushed ocean water outward to form a rim wave (blue) that was 4.5-km-high about 2.5 minutes after impact and traversed Earth's oceans.

Today's large tsunamis are simulated with shallow-water wave models, which assume that a tsunami's wavelength is much greater than the ocean depth. The water in the Gulf of Mexico, close to where the Chicxulub tsunami originated, deepens rapidly. The exact depth of the ocean 66 million years ago, however, is highly uncertain. But regardless of what the water depth may have been then, the model's results were about the same for all simulations, in which the depths ranged from 1000 to 3000 m. The tsunami for each simulation evolved into a 100-km-long wave—enough to satisfy the shallow-water assumption—10 minutes after impact.

Figure 2 shows the size of the waves that coastlines across the globe may have seen from the tsunami. Modeled tsunami amplitudes were largest near the impact site in the Gulf of Mexico, where coastlines were hit with waves more than 100 m high. Large-amplitude waves also inundated ocean basins far from the impact. The tsunami traveled at about 200 m/s in the open ocean, and it circled the globe in about two days. During that time, the tsunami approached most coastlines with an

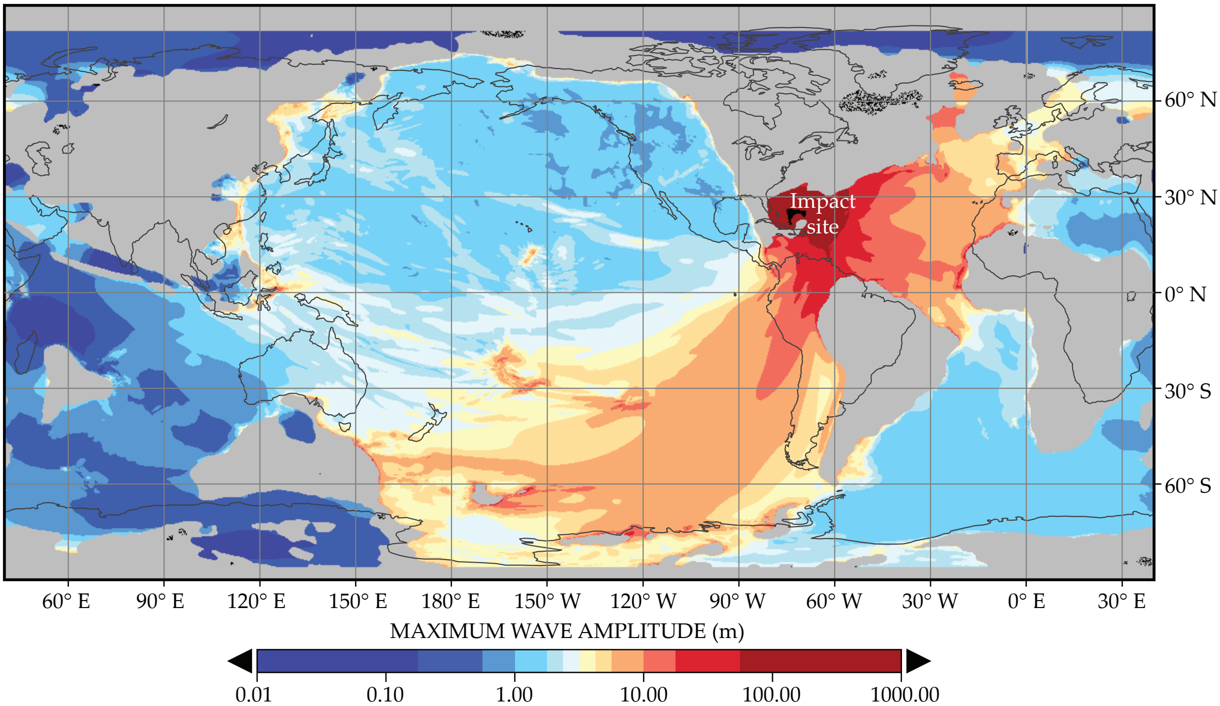


FIGURE 2. THE WAVE created by the dinosaur-killing asteroid impact made its way to most of the world's oceans. Colors indicate the maximum amplitude of the Chicxulub tsunami around the locations of ancient continents (gray), with the location of today's continents outlined in black for reference. Open-ocean tsunami amplitudes were highest in the Gulf of Mexico, the Caribbean Sea, and the North Atlantic and South Pacific Oceans.

amplitude of more than 1 m and hit many of them with an amplitude that exceeded 10 m. We expect that the Chicxulub megatsunami caused coastal flooding similar to that of the largest recorded tsunamis but on a global scale.

Because the Chicxulub impact took place 66 million years ago, Earth's continents and mid-ocean ridges were in different locations than they are today. To capture the tsunami's evolution, we used an ocean basin map with boundaries that were drawn by plate tectonics experts. The land bridge between North America and South America, for example, was absent, so the propagating tsunami wave easily entered the Pacific Ocean.

Connecting the geological record

According to figure 2, the tsunami wave height was at its largest in the Gulf of Mexico and the North Atlantic and South Pacific Oceans. Open-ocean tsunami velocities were as high as 100 m/s in the Gulf of Mexico and Caribbean Sea basins and lower yet significant farther away from the impact site. Flows that are faster than 20 cm/s are capable of eroding sediment, and our modeled open-ocean tsunami velocities exceeded that threshold in the North Atlantic, the South Pacific, the Gulf of Mexico, and the Caribbean.

To assess how realistic the simulation results may be, we compared them with sedimentary records located throughout the globe on land and in the open ocean. Open-ocean sediment cores are obtained with ocean drilling vessels that use the same technologies as oil drilling rigs. The deeper one looks into a sedimentary record, the further one travels back in time: The individual layers of rock can be dated through measurements of radioactive isotopes.

Consistent with the tsunami model results, sediment cores from 66 million years ago suggest that erosion was substantial in the Gulf of Mexico, the Caribbean, the North Atlantic, and the South Pacific. Sediments appear to be less eroded where the tsunami model predicts weak velocities, including in the Mediterranean Sea and the South Atlantic and North Pacific ocean basins.

We are already working on a follow-up study in which we simulate the coastal flooding caused by the tsunami. We are using estimates of land elevations and revised representations of continental geometries that take into account more geological data, especially near the impact site and in selected far-field regions that experienced large tsunami wave heights and high velocities. A new simulation that incorporates the detailed physics of coastal flooding should allow for a closer comparison between model results and geological records from the open ocean and coastal areas.

Brian Arbic thanks NSF and its Research Experiences for Undergraduates program for funding the work of Molly Range, an undergraduate student, on the project. Brandon Johnson thanks the developers of the iSALE shock-physics code. Vasily Titov acknowledges support for the work from NOAA, contribution number 5533.

Additional resources

►M. M. Range et al., "The Chicxulub impact produced a powerful global tsunami," *AGU Adv.* **3**, e2021AV000627 (2022).

►J. Erickson, "Dinosaur-killing asteroid triggered global tsunami that scoured seafloor thousands of miles from impact site," *Michigan News* (4 October 2022).

►C. G. Bardeen et al., "On transient climate change at the Cretaceous–Paleogene boundary due to atmospheric soot injections," *Proc. Natl. Acad. Sci. USA* **114**, E7415 (2017).

►J. V. Morgan et al., "The formation of peak rings in large impact craters," *Science* **354**, 878 (2016).

►H. J. Melosh et al., "Ignition of global wildfires at the Cretaceous/Tertiary boundary," *Nature* **343**, 251 (1990).

►L. W. Alvarez et al., "Extraterrestrial cause for the Cretaceous–Tertiary extinction," *Science* **208**, 1095 (1980). PT

Western Veil Nebula

At this time of year in the Northern Hemisphere, the Cygnus constellation, with its recognizable Northern Cross, is visible in the night sky. What's less visible is the Veil Nebula, a cloud of hot dust and ionized gas that's part of a large supernova remnant. The nebula formed about 10 000–20 000 years ago from the explosion of a massive star. Despite being roughly 2400 light-years away, the Veil has an angular size in the sky that's several times larger than the Moon. But even in ideal conditions, stargazers need a telescope to see it. And with increasing light pollution, astrophotographers are having a more difficult time capturing images of it.

This picture shows the western part of the Veil Nebula. Imran Sultan—a physics graduate student at Northwestern University—made the image, which was the 2023 astronomy winner of the Royal Society's annual photography competition. To construct the image, Sultan observed the sky over the Chicago suburbs for two nights and combined 52 exposures that each lasted for five minutes. "I was able to overcome the extreme light pollution," he says in the Royal Society's press release, "by using a special filter which only allows certain wavelengths of light to pass through." The emission that did get through comes from the bountiful hydrogen and oxygen gases in the Veil Nebula; they appear in the image as red and blue, respectively. (Image courtesy of Imran Sultan.) —AL

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