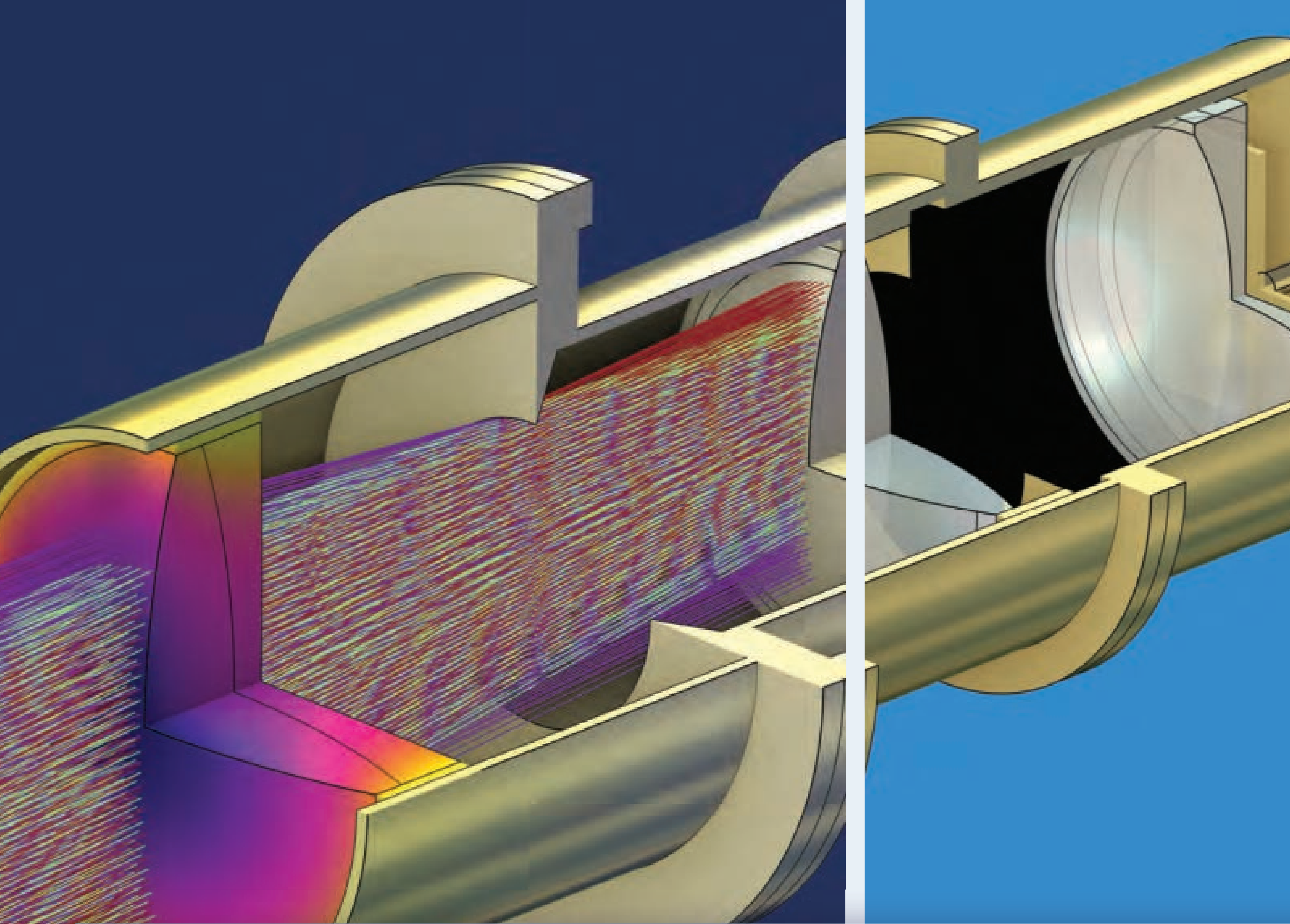


Shine Brighter in Optical Design

with COMSOL Multiphysics®

Multiphysics simulation drives the innovation of new light-based technologies and products. The power to build complete real-world models for accurate optical system simulations helps design engineers understand, predict, and optimize system performance.

» *comsol.com/feature/optics-innovation*

COMSOL

PHYSICS TODAY

February 2023 • volume 76, number 2

A publication of the American Institute of Physics

CLOGGING

The *JWST* reveals a nebula's neighbors

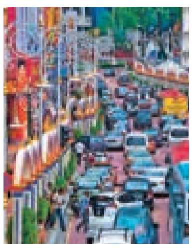
A Born–Oppenheimer approximation failure

Physicist-consultants

PHYSICS TODAY

February 2023 | volume 76 number 2

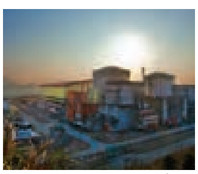
FEATURES



ON THE COVER: Clogging can take place whenever a suspension, made up of discrete objects in a fluid, flows through a confined space. The constriction may be microscopic, like the tiny pores of a filter, or macroscopic, like the narrow road that produced this traffic jam in Kuala Lumpur, Malaysia. To learn about how various clogs can happen and how to mitigate them, turn to the article by Brian Dincau, Emilie Dressaire, and Alban Sauret on **page 24**. (Photograph by Alexandre Roussel/Alamy Stock Photo.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org



ROY KALTSCHMIDT/LBNL

Antineutrino anomaly

Why do nuclear reactors seemingly emit fewer antineutrinos than theory predicts? That question has raised hopes that new physics, such as a sterile neutrino, is responsible. A recent study offers a compelling but more mundane explanation for the anomaly: flaws in the data underpinning theoretical models.

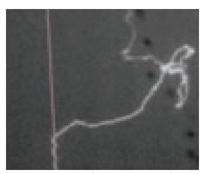
physicstoday.org/Feb2023a

THOMAS ANGUS/ICL

Q&A: Jessica Wade

"I'm going to be a professor who changes things," says materials scientist Jessica Wade. Through such activities as authoring almost 2000 Wikipedia articles about scientists from underrepresented groups, the Imperial College London research fellow is leading efforts to make physics more inclusive.

physicstoday.org/Feb2023b



HOUARD ET AL/*NAT PHOTON*/2023

Laser lightning rod

In the summer of 2021, four lightning strikes in the Swiss Alps were guided to the ground by a skyward-pointing laser. Directing lightning is the latest application of a serendipitously discovered technique that enables high-intensity laser pulses to create conducting channels in air.

physicstoday.org/Feb2023c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 76, number 2. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Periodicals postage paid at Huntington Station, NY, and at additional mailing offices. POSTMASTER: Send address changes to Physics Today, American Institute of Physics, 1305 Walt Whitman Rd, Suite 110, Melville, NY 11747-4300. Views expressed in Physics Today and on its website are those of the authors and not necessarily those of AIP or any of its member societies.

PHYSICS TODAY
www.physicstoday.org

DEPARTMENTS

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

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

PhysicsToday **@physicstoday**

AIP

SPS

SPECIAL OFFER

Post Your REU at SPS Jobs, FREE

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

Positions posted on SPS Jobs will also appear on:

- Physics Today Jobs
- American Association of Physics Teachers
- American Physical Society (APS Physics)
- AVS Science and Technology

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

Get started now!

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

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

Commentary

Astronomy from space after the *JWST*

Last year the *James Webb Space Telescope* (*JWST*) trained its sights on the distant universe and began returning stunning images in IR light that captured the imagination of the public around the world. The *JWST* has only just begun its exploration of the universe, yet astronomers are already looking to the future.

Every 10 years, panels of experts convened by the National Academies create a survey in astronomy and astrophysics to assess scientific frontiers and plan the capabilities that will keep advancing humankind's understanding of the cosmos. The *JWST*, originally called the *Next Generation Space Telescope*, was recommended as the top-priority large strategic mission in the 2000 decadal survey, and given its technical ambition and scale, it took more than two decades to be realized. But it was also conceived as costing less than 10% of the eventual price tag, in part because of wishful thinking by NASA and the scientific community and in part because the project began full-scale development before it was ready.

The 2010 decadal survey gave the highest priority for a large space mission to the *Nancy Grace Roman Space Telescope*, then known as the *Wide-Field Infrared Survey Telescope*, an observatory that will image large swaths of the sky in IR light to make three-dimensional cosmic maps of many millions of galaxies, elucidate the nature of the mysterious dark matter and dark energy, search for and image exoplanets, and explore many other topics in IR astrophysics. *Roman* is scheduled for launch in 2027, some 17 years after it was recommended as a priority. That delay was a result of funding issues directly related to the cost overruns on the *JWST*. Like the *JWST*, *Roman* was chosen to address the big, outstanding scientific questions identified at the time. But the delays bring with them costs in both taxpayer dollars and scientific opportunities.

The most recent astronomy and astrophysics survey, Astro2020 (for which the two of us served as committee cochairs), considered how to accomplish its charge to chart a course for the future of space astrophysics. In previous surveys, the committees have provided a rank-ordered list of concepts for large space missions, and in the decade that followed, NASA began development of the top project on the list as its next priority. Astro2020, however, took a different course.

STEPHAN'S QUINTET, as captured by the *James Webb Space Telescope*. The image, a mosaic constructed from almost 1000 separate IR images, reveals previously unseen details of the group of galaxies. (Courtesy of NASA, ESA, CSA, and STScI.)

The lessons of the past 20 years indicate that with business as usual, the next large mission beyond the *JWST* and *Roman* would be two or more decades away from its realization, and the next one after that more than four decades away. In the 1990s and early 2000s, NASA was able to launch a suite of Great Observatories—with four strategic missions, launched within a period of 13 years—that significantly advanced observational capabilities across a broad swath of the electromagnetic spectrum. The scientific impact of that panchromatic suite was enormous, as the universe could be simultaneously probed in many different and complementary ways, and it achieved leverage and insights well beyond what was imagined for any of the individual telescopes.

Astro2020 therefore includes a strategy to realize a new generation of Great Observatories: They have intervals between launches greatly reduced from the roughly two decades that is the recent norm yet still include capabilities at the scientific forefront. Far from business as usual, Astro2020 does not forward a ranked list of missions; rather, it imagines the Great Observatories Mission and Technology Maturation Program to study multiple mission concepts in the same decade. Individual mission cost targets would be appropriate for the scientific scope. While the survey prioritizes the first mission to enter the maturation program, it emphasizes that multiple missions should be studied this decade, so that if the one at the top of the list runs into problems, delays, or large cost overruns, backup options are ready.

First in the list of missions to be matured is an observatory that spans the wavelength range covered by the *Hubble Space Telescope* (which is 2.4 m in diameter and covers UV to near-IR) and has the collecting area of the *JWST* (6.5 m in diameter but mainly IR). The large IR/optical/UV (IR/O/UV) mission would have the ability to image a target planet while blocking out the light of its parent star, even when the star is 10 billion times brighter than the planet. It is an ambitious mission, on the scale of the *JWST*, yet the survey sets a target that it should not cost significantly more than the *JWST*.

The choice is motivated by the mission's ability to diagnose the atmospheres of planets outside the Milky Way to search for signatures of life—which, if detected, would change the way humans view their place in the universe. Like *Hubble*, the IR/O/UV mission would revolutionize our understanding of galaxies and stars and of the interstellar, circumgalactic, and intergalactic gases that give birth to them and would link them in a complex cosmic ecosystem. The IR/O/UV mission is technically challenging, and like the *JWST* and *Hubble* before it, it demands a major investment; NASA is the only agency worldwide capable of leading it.

Also compelling—and essential to advancing modern astrophysics—are a next-generation x-ray telescope and a mission sensitive in the far-IR. The former, with resolution matching that of NASA's *Chandra X-Ray Observatory* but with a vastly greater collecting area, would map hot, diffuse structures that are believed to feed the growth of galaxies and would peer back to find black holes forming in the early universe. The latter would unveil the dense regions of gas and dust enshrouding sites of star formation and the active central regions of many galaxies, and it would reveal the complex chemical processes that give rise to stars, planets, and ultimately life. With disciplined study and technology development, both missions can realize transformative capabilities on a size scale only one-third that of the large IR/O/UV mission. With strategic investment in the coming decade, both could also be ready to launch in quick succession.

Our "crystal ball" description of future missions and observatories beyond the *JWST* has focused on the largest space missions, but Astro2020 also recommends that NASA continue with a balanced portfolio of mission sizes from the large missions or Great Observatories described here down to probe, explorer, and smaller missions. Our committee was only tasked with planning future US-based activities, but in reality many of the projects will involve international partnerships, and implementation of the NASA road map will need to take into consideration missions led by the European Space Agency and other countries.

Astro2020 envisions a bright future, with eyes on the universe spanning the electromagnetic spectrum.

Fiona Harrison
(fiona@srl.caltech.edu)
California Institute of Technology
Pasadena
Robert Kennicutt
(rck@arizona.edu)
University of Arizona
Tucson
Texas A&M University
College Station

LETTERS

STEM volunteers

In response to Toni Feder's item "The US is in dire need of STEM teachers" (March 2022, page 25), I would like to make note of the work being done by the American Association for the Advancement of Science's STEM Volunteer Program (stemvolunteers.org), which I coordinate. We recruit STEM (science, technology, engineering, and mathematics) professionals to assist K–12 teachers in their classrooms.

The program began in 2004 and currently has 110 volunteers in four school districts in the Washington, DC, metropolitan area. Our retired volunteers commit to a few hours one day a week in the classroom for the school year. Those still working commit to a few hours every two to three weeks. Many volunteers exceed those commitments.

Volunteers help students learn subject matter through projects rather than by rote. They also present on technical subjects in the curriculum and organize "Ask a Scientist" sessions, in which they answer questions from groups of students.

One teacher wrote an email to a volunteer thanking him for the gift of his time and stating that it was an "absolute pleasure" to work with him. "You make science come alive for our children and I am very grateful," she wrote. "I will do all I can to encourage more schools to use the program and get visiting scientists."

Prospective volunteers are contacted through a variety of mechanisms—such as through societies' local sections, newsletters, the DC MIT Club, and retirement associations. The American Physical Society has supported the program from the beginning, including annually sending recruiting notices to its members in the DC metropolitan area.

I am convinced that a national program can be designed to produce a significant increase in the number of volunteers from the large number of STEM graduates. As of 2019, there were 12.3 million college graduates whose highest degree was in a science or engineering field, according to NSF's National Survey of College Graduates. A consortium of STEM societies is the best approach for implementing a program in support of K–12 STEM education.

Increasing the number of volunteers will not solve the teachers shortage, but volunteers can be a significant help, in particular with assisting teachers who have a limited background in STEM. And they can serve as substitute teachers, as several of our volunteers have done.

Don Rea
(donaldrea@aol.com)
AAAS STEM Volunteer Program
Washington, DC

Mind and matter

Andrew Zangwill's March 2022 article (page 28) presents an insightful portrait of Philip Anderson in dynamic,

human terms. I was particularly drawn in by Zangwill's mention of Anderson's Anglophilia and his association with leading researchers at the University of Cambridge. In that context, two of Anderson's "four facts"—that computers will not replace scientists and that good science has aesthetic qualities—resonate with Brian Josephson's interests in the past 20-odd years.

I met Josephson at an international conference, titled Home and the World: Rabindranath Tagore at the End of the Millennium, which was held by the University of Connecticut in September 1998. Josephson spoke about the poet-philosopher Tagore (1861–1941) and science.[1] From my relatively brief encounter with him, I understood at the time that Josephson was especially interested in the area of mind–matter interactions, and that, of course, had some relevance to the well-known 1930 conversation that Tagore had with Albert Einstein on reality and the human mind.[2] Mind–matter interactions have also been an area of sustained interest for many leading scientists, including Ilya Prigogine and Roger Penrose.

It is also quite noteworthy that Zangwill mentions Charles Kittel as one of Anderson's mentors at Bell Labs. Many of us pursuing physics and engineering in India in the 1970s were introduced to Kittel's classic textbook *Introduction to Solid State Physics*, which was foundational to our understanding of the subject.

References

1. B. Josephson, in *Rabindranath Tagore: Universality and Tradition*, P. C. Hogan, L. Pandit, eds., Fairleigh Dickinson U. Press (2003), p. 107.
2. M. Popova, "When Einstein met Tagore: A remarkable meeting of minds on the edge of science and spirituality," *Marginalian* (27 April 2012).

Monish R. Chatterjee
(mchatterjee1@udayton.edu)
University of Dayton
Dayton, Ohio

CO_2 air-capture costs

David Kramer's "A windfall for US carbon capture and storage" (January 2022, page 22) mentions the $3.5 billion appropriated by the US government for direct air capture. I would like to point out that the energy costs of capturing carbon dioxide already diluted in the atmosphere would be prohibitive.

Methods tried so far employ a reusable absorber cycled between absorption and emission, with an input of energy required at one or both parts of the cycle. The unavoidable energy requirement for a cycle can be calculated from the entropy change ΔS of the CO_2 going from its present atmospheric concentration of about 400 ppm to a concentration needed for disposal or use, say 1 atmosphere.

Per unit mass and at room temperature T, that energy would be $T\Delta S = RT/M \ln(10^6/400) = 4.4 \times 10^5$ kJ/ton (t), where R is the molar gas constant and M the molar mass. If you assume the energy is applied electrically, and at a present US price of 12¢/kWh, the energy cost is \$15/t. So far there are no reports of technologies that are anywhere close to that energy requirement or cost.

Earth's atmosphere weighs 5.2×10^{15} t. The unavoidable entropy cost to remove just 1 ppm (by volume) of CO_2, or 7.9×10^9 t, would be \$120 billion. After recovery at 1 atmosphere, there are the added costs of disposal, which is complicated by the residual atmospheric gases in the recovered CO_2.

The cost could be reduced if the energy is somehow supplied directly rather than after conversion to electricity. But no energy source is free because its energy could otherwise be converted to electricity and sold.

The costs of mineralization are more difficult to estimate. The absorber is used only once, not cycled. Costs might include those for accessing, processing by crushing and dispersing, and gathering and disposing of the absorber.

John Tanner
(pust@datawav.net)
Idaho Falls, Idaho PT

CONTACT PHYSICS TODAY

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

Neighboring stars shaped a planetary nebula

Early images from the *JWST* reveal previously unseen fingerprints of stellar interactions with a dying star.

On 12 July 2022, NASA released the first images from the *James Web Space Telescope* (*JWST*). Chosen to demonstrate the telescope's capabilities, the five results included a distant galaxy cluster, an atmospheric spectrum of a distant planet, a quintet of galaxies, a stellar nursery, and the one Orsola De Marco was most excited about: NGC 3132, a planetary nebula (PN) informally called the Southern Ring Nebula. De Marco, of Macquarie University in Sydney, Australia, stayed awake late to watch the announcement broadcast from Maryland. As soon as she saw the new PN image, she noticed something unexpected, and before the announcement was over, De Marco and her colleagues in the PN research community were exchanging emails about the puzzling result.

A PN is the end stage of an intermediate-mass star, and such nebulae are the biggest carbon and nitrogen producers in the universe. The glowing shells of ionized gas form when a star of 1–8 solar masses dies. A dying star expands, and the outermost material escapes more easily. The star eventually loses enough mass to reveal its hot core, which, no longer burning, becomes a white dwarf. The star temporarily emits UV radiation, which ionizes the cast-off gas and creates the characteristic glow. The PN stops glowing once the star dims too much to emit enough UV radiation.

The Southern Ring Nebula, which is about 2500 light-years from Earth, formed from a dying star of about 3 solar masses. It had already been imaged by the *Hubble Space Telescope* and various ground-based observatories. The *JWST* image in the near-IR, shown in figure 1a, resembles previous observations, which have largely been in the visible and near-IR. The initial surprise came from the *JWST*'s mid-IR image, seen in figure 1b. A red dot near the center of the nebula shows the dying star brightly shining in the IR. The star is still quite hot—around 130 000 K. At that temperature, it should emit blue or UV light. Where was the IR coming from?

De Marco and 68 colleagues from around the world figured out the answer to that question and the origins of several other curious features of the Southern Ring Nebula.[1] The astrophysics detectives identified evidence of unseen stars and pieced together how those stars shaped the nebula. Their study contributes to an open question about how PNs form and evolve.

Dusting for fingerprints

PNs have nothing to do with planets. The "planetary" in the name is a holdover from the 18th century when William Herschel noted that a PN's round shape and uniform color—as seen through telescopes of the time, anyway—resembled a planet. Since then, improved observational tools have revealed that PNs are far from uniform and take various shapes. Around 10–20% have the spherical form that would be intuitively predicted for a spherical star expelling a spherical wind.[2]

Over the past 20 years, researchers have suspected that to get the nonspherical PN shapes seen in the sky, the dying star must interact with a neighboring star. They have looked for evidence of binary stars at the heart of PNs, but directly observing them can be difficult when the companion star is too dim or too close to the brighter central one to be detected. Indirect signatures, however, have suggested that a large fraction of PNs have close companion stars.

Astronomers already knew that the Southern Ring Nebula was at least a binary star system. The PN's bright star in near-IR (see figure 1a) and visible images is not the one that formed the PN, even though it appears near the center, but rather a companion star threefold the

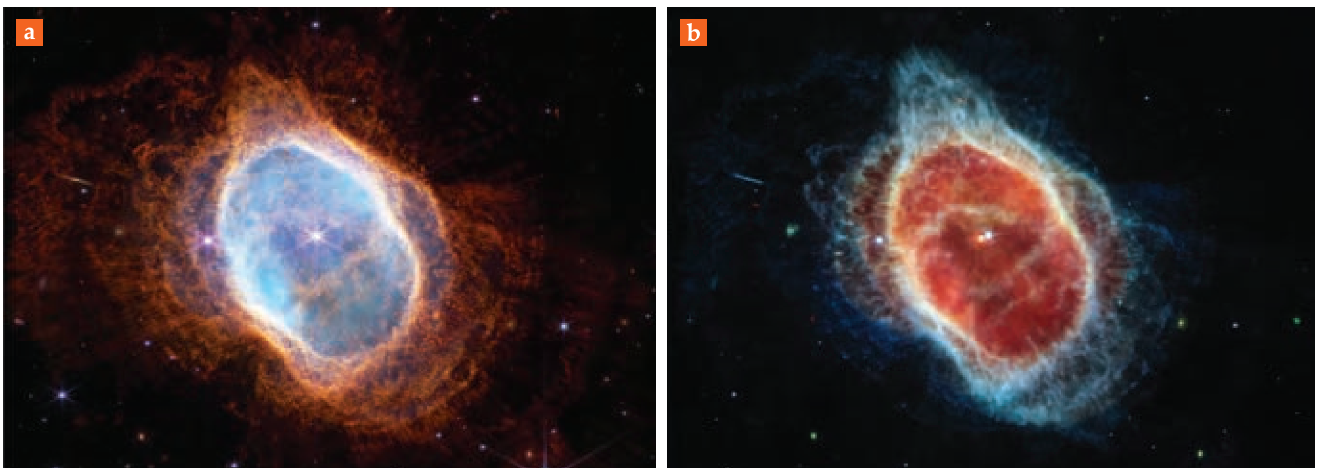


FIGURE 1. PLANETARY NEBULAE are glowing shells of gas that form around intermediate-mass dying stars. The one shown here—NGC 3132, or the Southern Ring Nebula—was among the early objects imaged by the *James Webb Space Telescope*. **(a)** The near-IR image resembles earlier ones taken by other telescopes, although the increased light collection reveals a pattern of incomplete concentric rings (most noticeable in the lower right quadrant) in the outermost gas. **(b)** The mid-IR image surprised astronomers because the dying star, the red center dot, wasn't expected to emit in the IR. (Courtesy of NASA/ESA/CSA/STScI.)

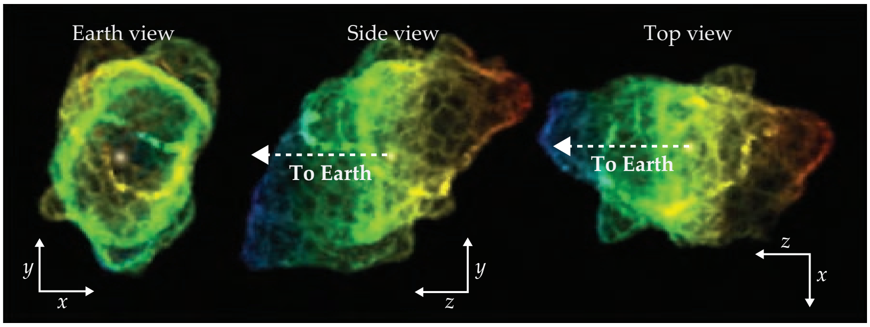


FIGURE 2. A RECONSTRUCTION of the Southern Ring Nebula's ionized gas bubble, shown here at different orientations, clarifies the nebula's structure and how it relates to stellar interactions. The roughly egg-shaped nebula features bumps that come in pairs on opposites sides. From simulations, researchers propose that the bumps are from pairs of plasma jets ejected by two stars near the dying star. From the bumps and other evidence, they deduce that the nebula's center hosts four or five stars. The colors here show the Doppler shift of the light that reaches Earth. (Adapted from ref. 1.)

mass of the Sun. In 1999 *Hubble* just made out the PN's faint dying star next to the bright companion. The two stars are too far apart to interact in any substantial way, however, so that visible companion couldn't be responsible for the shape of the nebula. Or for the unexpected IR emission.

"When something is shining in the IR, it has to be relatively cool, and in this case, it has to be dust," says De Marco. Dust around a central star warms just enough to emit in the mid-IR or burns away. Typically a hot star at a PN center ends up surrounded by a dusty disk because of a strong interaction with a nearby star.

The researchers concluded that at some point, the dying star that made the Southern Ring Nebula had a close neighbor strip away some of its mass and leave behind dust. As a result, in the mid-IR image shown in figure 1b, the dying star shows up as a distinct red dot near the PN's center, with its previously known distant companion appearing as the bluish-white dot next to it. The close dust-producing companion may still be there, but it's too dim to see, or it may have merged with the dying star.

Bump in the night

The *JWST*'s look into the mid-IR had thus revealed unexpected dust. "IR observations tend to be a little rarer" than those in the visible, says De Marco. "Whenever you open that window, there are always surprises." The dust was the first of several such surprises enabled by the *JWST*. As the largest space telescope, it captures more photons and picks up fainter features than previous instruments. One such feature is the pattern, albeit weathered and incomplete, of concentric rings formed by the outermost gas of the Southern Ring Nebula, just visible in figure 1a. "In planetary nebulae, there's a long history of studying rings like those," says De Marco. "The best explanation is that when the nebulae form and slowly puff out, the gas goes past whatever is in the way. And if the dying star has a companion that goes around it, the companion stirs the gas and creates an imprint of a spiral." As the spiral expands, it forms rings.

Researchers already knew how quickly a PN's gas expands, usually around 20 km/s. So Shazrene Mohamed of the University of Cape Town, South Africa, modeled how far the stirring star is from the dying one through the distance between rings—effectively the companion star's orbital period. She found a distance of 40–60 AU, which is farther than the neighbor responsible for the dust and much closer than the visible companion. So the tally increased to four stars: the dying one; the distant, previously known one; the close, dust-producing one; and now a middle-distance, ring-producing one. "We've got a quartet," says De Marco, "and though indirect, the evidence is pretty strong" because it's based on long-standing models for nebulae and binary interactions.

More-tenuous evidence points to a possible fifth star. A three-dimensional reconstruction of the PN ionized gas bubble, shown in figure 2, was generated by Wolfgang Steffen and his company, Ilumbra, which models astronomical objects. The researchers 3D-printed the final fit. As they held the object, which was roughly egg-shaped with many bumps, they noticed the bumps came in pairs on opposite sides, and the lines between the pairs went roughly through the center star. A likely way to produce such bumps is plasma jets, which come in pairs, that originate near the dying star. The number and diverse orientations of the jets required to account for the PN bumps are possible only from interactions between three stars. So in addition to the close companion responsible for the dust, there may be another star that's closer than the ring-producing or visible companion. A five-star system is unusual but not impossible. For stars the mass of the dying one, models predict around 10% have four companions.

The game is afoot

The researchers compare their work to a murder investigation: A neighboring star that contributes to and expedites the central star's death murders it, effectively. To do so, the companion needs to be close enough to interact strongly and pull away mass. The visible companion is thus an innocent bystander, and the one that made the rings is an accomplice but didn't strike the lethal blow. The one or two closest companions are the murderers.

The investigation hinged on the wavelength range, improved resolution, and greater light collection provided by the *JWST*. (See "First observations test *JWST* capabilities," Physics Today online, 17 November 2022.) "Indirect evidence is always difficult to deal with, and if it's indirect and fuzzy, it's even harder," says De Marco. "But when the indirect evidence is sharp, it becomes useful." In the future, similar analyses could be done on additional PNs with varied shapes—for example, the Cat's Eye Nebula, whose name evokes its complicated and unusual structure.

"With the Southern Ring Nebula, we have a superb example of a complex form that is not very different from many other nonspherical planetary nebulae," says George Jacoby of NOIRLab. "But now with sufficient data to derive an excellent set of models and explanations for those complexities, we can be much more confident generalizing the physics of those planetary nebulae."

Heather M. Hill

References

1. O. De Marco et al., *Nat. Astron.* **6**, 1421 (2022).
2. G. H. Jacoby et al., *Publ. Astron. Soc. Aust.* **27**, 156 (2010).

A solid-state failure of the Born–Oppenheimer approximation

According to a keystone principle of molecular physics, atoms striking semiconductor surfaces shouldn't excite surface electrons. But they do.

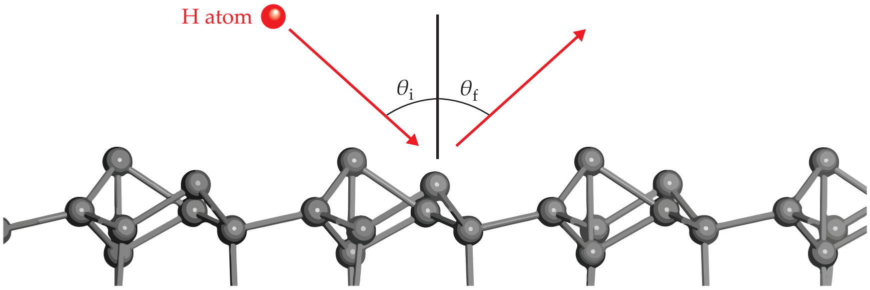


FIGURE 1. TO TEST THEORY against experiment in surface chemistry, researchers study a simple atom–surface interaction: bouncing hydrogen atoms against a solid surface at a specified speed and incidence angle θ_i, and measuring the speed and angle θ_f of the scattered atoms. The surface shown here is the (111) crystal surface of germanium in the reconstruction denoted by $c(2 \times 8)$. (Courtesy of Kerstin Krüger.)

Electrons are light and quick; atomic nuclei, comparatively, are heavy and slow. The gap in mass and time scales means that electronic and nuclear dynamics can usually be considered separately, as Max Born and J. Robert Oppenheimer pointed out in a 1927 paper[1] that has heavily influenced how chemical physicists think about atoms and molecules to this day: Electrons gracefully flit around ponderous nuclei and instantaneously follow them wherever they go.

The intuitive Born–Oppenheimer picture doesn't always hold, but it generally does a pretty good job. And in situations where it does break down, it's usually clear what's going on. In some chemical reactions, for example, the nuclei are known to rearrange too rapidly for the electrons to keep up (see PHYSICS TODAY, August 2021, page 14).

Now, however, Kerstin Krüger of Georg-August University in Göttingen, Germany; her PhD advisers, Oliver Bünermann and Alec Wodtke; and their colleagues have found that in at least some interactions between atoms and solid surfaces, the Born–Oppenheimer approximation isn't just a little bit off—it fails dramatically. And they're not quite sure why.

The researchers fired hydrogen atoms at the surface of germanium, a semiconductor, and found that a large fraction of the atoms were losing a significant portion of their energy.[2] A close look at the data suggested that the atoms were whacking Ge surface electrons hard enough to excite them from the valence band into the conduction band. Under the Born–Oppenheimer approximation, that's not supposed to happen. The electrons are supposed to be nimble enough to get out of the way.

Testing theory

If not for the Born–Oppenheimer approximation, computational chemistry might never have gotten off the ground. To simulate the dynamics of a molecule, researchers would need to solve the multiparticle time-dependent Schrödinger equation to compute the dynamics of all of the nuclei and electrons simultaneously.

Thankfully, they don't have to do that. Instead, they can consider the nuclei to be moving around in a potential-energy landscape created by the electrons. The means of estimating that potential have grown over the years in sophistication and accuracy (see PHYSICS TODAY, December 2013, page 13), from empirical interatomic potential-energy functions to fully quantum calculations of the electrons' actual energy. What they have in common, though, is the assumption that the electrons can keep up with the nuclei: The potential energy is a function of where the nuclei are, not where they've been or where they're going.

Molecular-dynamics simulations have yielded many valuable insights into microscopic molecular motions on a level that's not accessible to experimental methods. But it's important to keep testing their results against experiments in order to make sure they're giving the right answers and that their assumptions are warranted.

Several years ago, Bünermann and colleagues observed that there was a gap in researchers' understanding of how well molecular dynamics simulations apply to chemical reactions at solid surfaces. Surface reactions are widely used in industrial chemistry, to profound effect on humanity (see PHYSICS TODAY, September 2018, page 17). But they're famously messy, and the challenge in testing experiments against theory is to find surface processes that are both experimentally feasible and theoretically tractable.

Most surface reactions involve polyatomic molecules, whose many degrees of freedom make them complicated to model. On the other hand, single reactive atoms (that is, not noble gas atoms) can be experimentally challenging: They can't just be pulled out of a bottle, so researchers need to make them on the fly by breaking apart molecules.

The experimental apparatus for dissecting the simple atom–surface interaction shown in figure 1—shooting H atoms at a surface with a specified speed and direction and measuring the speed and direction at which they bounce back—required a specialized combination of components, and it had never been built before. So Bünermann and colleagues built it.[3] And they've been using it to study H-atom scattering off all sorts of surfaces.

Mysterious excitations

The apparatus was designed to hunt for violations of the Born–Oppenheimer approximation. "The approximation has always been on our mind," says Bünermann, "because H atoms sometimes stick to surfaces. So there must be a violation somewhere, because otherwise all the atoms would just bounce back."

When the researchers fired H atoms at a metal surface, for example, they observed a minor deviation from the Born–Oppenheimer approximation that they could account for with a perturbative correction: Electrons at the surface be-

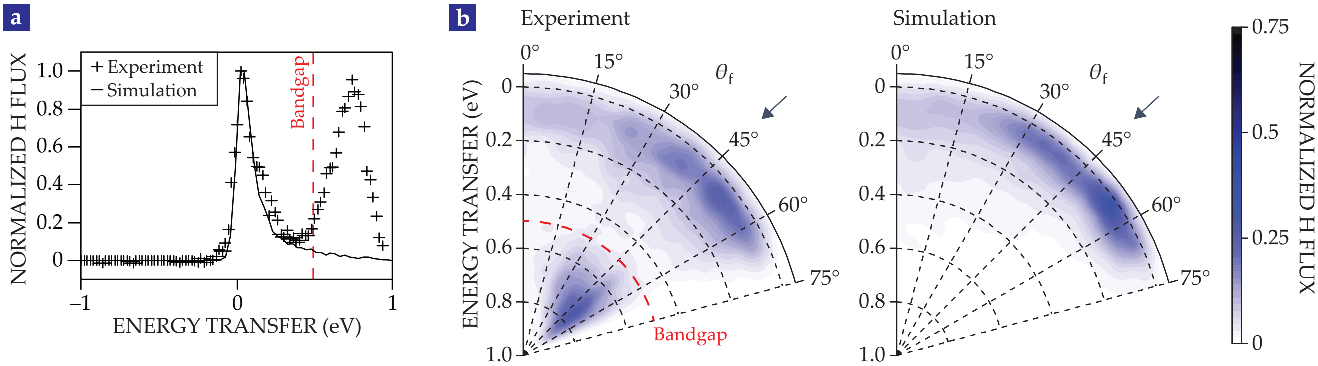


FIGURE 2. THE ENERGY DISTRIBUTIONS of hydrogen atoms scattering off a germanium surface reveal an unexpected peak. **(a)** Atoms with an initial energy of 0.99 eV and an initial angle of 45° emerge in a bimodal distribution of energies: Some atoms lose just a little energy to the surface, while others lose a lot. The first peak is quantitatively accounted for in molecular-dynamics simulations, but the second peak is completely absent. **(b)** Data for the same experiment and simulation are resolved according to the scattering angle θ_f. The dashed red line is the Ge surface bandgap. (Adapted from ref. 2.)

haved like a viscous fluid that exerted a drag force on the intruding H atoms.[4] As a result, atoms bouncing off the surface emerged with less energy than they otherwise would have. Some of them lost all their energy, and they remained on the surface.

In the new experiments on semiconductor surfaces, however, the electron-drag theory doesn't apply: Electrons can't act like a viscous fluid when they're locked into a regimented set of valence-band quantum states. So Bünermann and colleagues expected the Born–Oppenheimer approximation to hold—but that's not what they observed. As shown in figure 2a, the scattered atoms have a bimodal energy distribution. Some atoms lose just a little of their energy when they bounce off the surface. Others lose a lot.

The first of those peaks is quantitatively reproduced by the molecular-dynamics simulations performed by Bünermann's collaborators Yingqi Wang and her adviser, Hua Guo, at the University of New Mexico, and it's fully consistent with the Born–Oppenheimer approximation. But the simulations don't reproduce the second peak at all.

One way that atoms could lose so much energy is if they stick to the surface, thermalize, and desorb. But from a look at the atoms' angular distributions, Bünermann and colleagues are pretty sure that that's not what's happening. As shown in figure 2b, atoms in both channels scatter at angles θ_f close to the incidence angle $\theta_i = 45°$. If the atoms were thermalizing and desorbing, one would expect a much broader angular distribution independent of θ_i.

The researchers have two reasons for concluding that the mysterious peak is due to H atoms exciting Ge electrons from the valence band into the conduction band. First, in all their experiments, the peak sets in just above Ge's surface bandgap of 0.49 eV. Second, when they increase the H atoms' initial energy, the peak gets bigger. In the experiments shown in figure 2, about half of the atoms with an initial energy of 0.99 eV emerge in the unexplained high-energy-transfer channel. But for an initial energy of 6.17 eV, that portion jumps to more than 90%. If the atoms excite electrons because they're moving too fast for the electrons to get out of the way, it stands to reason that speeding up the atoms would increase the magnitude of the effect—just what the researchers see.

Beneath the surface

Bünermann and colleagues' published experiments focus on the Ge(111) surface, because it's the easiest semiconductor surface for Wang and Guo to tackle in their simulations. Crystal surfaces tend to undergo reconstruction—that is, their atoms rearrange into structures that differ from the bulk—and the Ge(111) reconstruction, denoted by "$c(2 \times 8)$," is well understood. But they've also performed experiments on several other semiconductor surfaces, including Ge(100) and the (111) and (100) surfaces of silicon. And they see essentially the same effect each time.

Although they don't yet have a theory that can quantitatively explain their observations, they're starting to understand why the atoms separate into two peaks. Because of the reconstruction, the Ge atoms at the surface aren't all chemically equivalent. "We're pretty sure that the branching is connected to which one of the different germanium environments the atom hits," says Bünermann.

The researchers are focused on the basic-research task of understanding the fundamental role of the Born–Oppenheimer approximation in surface chemistry. But they also have some ideas for how the phenomenon they've discovered might be applied. Because atom–semiconductor collisions promote electrons into the conduction band with surprising efficiency, the excited electrons might be detected electrically, which could be the basis for a new type of sensor.

The scattering experiment might also be adapted into a new technique for studying the electronic structure of surfaces. Plenty of techniques for looking at surfaces are already available, such as photoelectron spectroscopy and electron energy-loss spectroscopy, but their signals are sometimes influenced by bulk electrons too. "Here, we're sure that we're seeing only surface states," says Bünermann, "and those are normally not so easy to access."

Johanna Miller

References

1. M. Born, R. Oppenheimer, *Ann. Phys.* **389**, 0457 (1927).
2. K. Krüger et al., *Nat. Chem.* (2022), doi:10.1038/s41557-022-01085-x.
3. O. Bünermann et al., *Rev. Sci. Instrum.* **89**, 094101 (2018).
4. O. Bünermann et al., *Science* **350**, 1346 (2015).

PT

Consulting can be stimulating and lucrative for physicists

Skills from physics training transfer well to a career that comes in many flavors.

"If you need someone to tell you what to do, consulting is not for you," says Stephanie Chasteen, a condensed-matter physicist turned physics education specialist who makes a living with her one-woman physics education consulting business. On the other hand, for physicists who are self-motivated, organized, and communicate well, consulting may be an attractive option at any career stage—from freshly minted bachelors or PhDs to mid- or late-career professionals who want to try something new.

For early-career scientists, consulting firms can be a good fit. As a fifth-year graduate student at the University of Michigan studying magneto-optical phenomena with ultrafast lasers, Elizabeth Dreyer realized she didn't want to spend the rest of her life "alone in a dark basement." After earning her engineering PhD in 2018, she joined Boston Consulting Group. It is one of the Big Three strategy consulting firms, along with McKinsey & Company and Bain & Company; each employs thousands of consultants. Now, after four and a half years, she is moving on. Over the same period she might have spent as a postdoc, she says, "I learned more than an MBA's worth of knowledge and got paid four times as much." By the end she was earning more than $200 000 a year. The network she has built through consulting and the assistance the company provides its alumni, she adds, give her confidence about finding her next job.

Another mode of consulting is the side gig, which some active and retired faculty and research lab scientists engage in. Such work can supplement income or serve as a low-risk transition to a full-time consulting career.

Fast paced and project based

At consulting firms, consultants regroup for each project. A project can last days to years, but about 12 weeks is common. Consultants dive deep into questions for their clients: Is the timing right for an automaker to move into the electric-vehicle market? What's the best way to market a particular drug? What types of investments in digital transformation would best benefit an oil company? How can a manufacturer produce steel more efficiently?

BRIAN MOORE

INDEPENDENT CONSULTANT Stephanie Chasteen says her work is intellectually stimulating and impactful, lets her "geek out on statistics," and allows for work–life balance.

Consultants spend their time analyzing data, interviewing people, participating in meetings, formulating strategies, and putting together presentations. They may also implement recommendations, seek projects and clients, and engage in professional development. At the big firms, the culture is often "up or out," and after a couple of years consultants who stay are promoted to managers. Before the pandemic, traveling to a client's site was the norm; consultants were often away from home Monday through Thursday. Now some are traveling again, but remote meetings have become a staple.

After earning his PhD in condensed-matter physics at Harvard University in 2014, Michael Yee joined McKinsey and stayed eight years. At the firm, "to work in condensed-matter stuff, I'd have had to travel to the West Coast or internationally," he says. Instead, to stay near his home in Toronto, he chose to focus on banking and insurance. He worked on strategy and business transformation. "My teams would perform complex analytics on the market opportunity. I would help clients hire people. For one client, we launched a new business brand."

WOODY CARRICK

JOHN BARENTINE measures the surface brightness of a streetlight with a luminance meter. He saw an opportunity to consult for parks and municipalities that want to improve their lighting or be accredited as a dark-skies location.

While working as a consultant, Yee says, "I used my creative brain and my analytical muscles. I found it exciting." And he felt good about the social impact. Some of the products his clients were rolling out were aimed at people in under-served markets who didn't have access to typical banking or lending products. This past September, Yee left his consulting position to become vice president of financial services for Canada Post because he "wanted to run a business rather than just advise."

Maggie Seeds joined Clarkston Consulting, a North Carolina–based firm with about 300 consultants, after earning her bachelor's in physics at Appalachian State University in 2012. "You hear that physics degrees open doors, but I didn't know what doors," she says. She was drawn to consulting when she saw, through the interview process, that she could get results faster and see the impact faster than in research. She first completed a boot camp that the firm offered. "It was like consulting 101," she says, "and we were paid."

"With consulting, you never know what type of project you'll be on next," says Seeds. At Clarkston, the teams for each project are put together by an in-house group, so the individual consultants don't market themselves internally like at some larger consulting firms. Most of her projects have involved helping companies implement and customize software. But a few years ago, Seeds and two colleagues started an analytics team within the firm. They work with clients on projects related to data storage, data quality, visualization of historical data, and predictive algorithms, and they create apps for clients to integrate predictions into their businesses.

In shifting into data science, Seeds says, "I was looking for a new professional challenge. I like being part of the decision making on what to build and then building it." One tool Seeds made uses artificial intelligence to recommend the respective quantities of black beans and pinto beans that her client should stock. The forecast was based on such input as the neighborhood, season, and previous sales.

As a graduate student working on the thermodynamics of single molecules at North Carolina State University, Zubair Azad was on the path to becoming a professor. But after stints as a postdoc and as a Fulbright scholar in Barcelona, Spain, he wound up at Deloitte in 2019. The appeal of consulting, he says, included the possibility to make a difference through advising decision makers, implementing recommendations, and learning new things with each project.

Early on Azad worked on digital issues in R&D, manufacturing, and supply chains in order to help companies commercialize products. He has since pivoted to establishing public–private partnerships that address societal needs. In his consulting role, he was recently named a fellow of the World Economic Forum, where he is involved with developing a framework to provide equitable access to vaccines through regionalized vaccine manufacturing. "Consulting is whatever you make of it," he says. "You can re-engineer your career as your goals and ambitions shift over time."

Sometimes, Azad says, he uses his technical skills and physics know-how in his consulting. And his PhD gives him "instant credibility" with clients and colleagues. But the softer skills are essential. The ability to communicate is "going to make or break your career," he says. "How are you with people management? Are you a good leader? Can you forge consensus?" That most important skill for consulting is the one that is least addressed during the PhD, he says.

Transferable skills

Although it can be hard to see from a researcher's vantage, scientific training builds many of the skills needed to be successful as a consultant, says Manu Lakshmanan. He worked at McKinsey for a year and a half after earning his PhD at Duke University in biomedical engineering in 2015 and doing a postdoc at the National Institutes of Health. Physicists have experience developing quantitative models and making approximations, he says. Knowing how to search the literature and communicate in writing are also transferable from academia, Lakshmanan says. And presenting talks and posters at conferences is not that far from persuading colleagues to work together on a project or convincing clients to consider implementing recommendations.

Among the skills that he had to hone when he entered consulting, Lakshmanan says, were being comfortable with qualitative data, working in teams, marketing himself within the firm, and explaining results on an impromptu basis. For physicists thinking of moving into consulting, he says, "extroverts can be more confident that it's a good move. I'm an introvert. It was a challenge."

Some people find the merry-go-round of colleagues and managers frustrating.

Still, many physicist-consultants say their colleagues and clients are a favorite part of their jobs. "The people I get to work with are not only smart and capable but kind and compassionate," says Azad. "I now care more about who I work with than what I do." Less attractive are the long hours—80-hour work weeks are common. "Work–life balance is a perennial problem," he says.

Independent consultants

By contrast, independent consultants point to work–life balance as a perk. Chasteen's education consulting largely involves her serving as an external evaluator; many grants and proposals require such oversight. She helps get projects underway, collects data, and provides feedback, mostly in the areas of departmental overhaul and faculty uptake of innovative teaching techniques. For example, in an analysis of PhysTEC, which promotes K–12 teaching as a career for physics graduates, she found that the program would benefit by expanding its focus on master's recipients.

Chasteen's consulting projects extend over years, which makes it possible for her to organize her time and maintain a consistent income. She chooses how much to work—typically 35–40 hours a week. She also chooses who to take on as clients.

"I don't work for people who just want to check a box," Chasteen says. "I work for clients like the American Physical Society and the American Association of Physics Teachers, who take the data I provide to make improvements." She started her business in 2009 after more than a decade of immersing herself in physics education issues. Having expertise in a niche area is typical for physicists who make a go of independent consulting.

Astronomer John Barentine launched his consulting business in summer 2021. From years working at the International Dark-Sky Association, a nongovernmental organization, he recognized a need he was well-suited to fill: advising parks, nature reserves, and municipalities on how to improve their lighting. "If they receive accreditation for dark skies, it's a badge of pride and it drives tourism," he says. "Some clients just want advice. Others want help writing bylaws or deeper involvement in the implementation." Barentine says he's had months where he brought in more than enough to cover expenses and had months with no work. It's too soon to see whether he'll succeed long term, he says.

Scot Kleinman is also new to independent consulting. He left his job as associate director of the Gemini Observatory last spring to go solo. "I was feeling confined in a normal job environment," he says, "and I was not having opportunities to work in the breadth of what I could offer." So far, the demand in astronomy is keeping him busy. "Sometimes it's a matter of making connections for a client. Sometimes they need fresh eyes." He does technical writing, develops materials for funding proposals, and has stepped in as interim manager for an instrument under construction for the Thirty Meter Telescope.

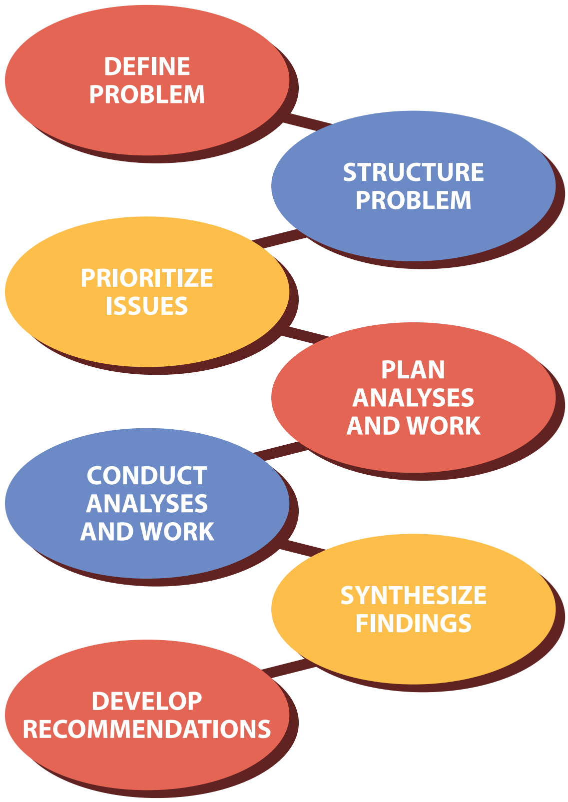
Kleinman says he finds the work more meaningful than he expected. "I get paid for doing interesting things that I'm good at and that people care about. Nobody hires a consultant to do work that isn't valued."

After 24 years at NIST, Carl Williams retired in late 2021 and launched CJW Quantum Consulting. "I wanted to more strongly engage and help make the science transition to applications than I could from within the government," he says. His clients include professional societies and companies, which he advises on market strategies, policy, and export controls related to quantum technologies.

Given that he already had a pension plan, Williams wasn't worried about money, but he says he's making plenty. Other former government scientists who have become consultants told PHYSICS TODAY that they bring home incomes in the several-hundred-thousand-dollar range. (See the article by Williams, PHYSICS TODAY, October 2022, page 50.)

As major pluses, independent consultants point to the flexibility of their schedules, being their own boss, and the impact they can make. Chasteen also likes that she can take the time to immerse herself in a topic. "My faculty colleagues are spread thin and juggling so much," she says. "I can dig much deeper into one thing at a time, and that's satisfying."

On the downside, many of them dislike the bookkeeping and rate-setting parts of their responsibilities and having to arrange for health insurance, social security, and retirement. Time management and negotiating skills are essential, they say.



THE SEVEN-STEP PROBLEM-SOLVING strategy at McKinsey & Company bears resemblance to the scientific method. (Based on the 13 September 2019 episode of the McKinsey Podcast.)

Side hustles

Daniel Lathrop, a physics professor at the University of Maryland, is occasionally approached by companies for help in nonlinear dynamics and neuromorphic computing. University policies on consulting vary; his employer permits faculty to devote up to a day a week to consulting. It's important to be transparent with the university and to avoid conflicts of interest, Lathrop says. Consulting "even a couple of hours a week is a substantial supplement to income for a midcareer academic professional," he says.

Laura McCullough, a physics professor at the University of Wisconsin–Stout, also consults on the side. But she's looking to build up consulting as a possible full-time career. She evaluates departmental climate from a DEI (diversity, equity, and inclusivity) point of view. In that role, she conducts climate surveys and site visits and reviews departmental bylaws. She tailors surveys to a given department's request. "I was doing occasional gigs and discovered I really liked it," she says.

McCullough is motivated to expand into consulting because resources are low at her state school. And workloads have grown, she says. She is a full professor who has been at her institution for 22 years and makes $80 000. "I love my job," she says, "but looking ahead, consulting may ease retirement."

"Right now it's a side hustle," McCullough says. "I don't think I could make a living at this yet. I'm still building up." But DEI has been a passion for 25 years. "I love helping organizations and departments make life better."

Toni Feder PT

CLOGGING:
The self-sabotage of suspensions

ALAMY.COM/RONALD RAMPSCH

Brian Dincau is a doctoral student, **Emilie Dressaire** is an assistant professor, and **Alban Sauret** is an associate professor, all in the department of mechanical engineering at the University of California, Santa Barbara.

Brian Dincau, Emilie Dressaire, and Alban Sauret

Whether it's pipes, highways, or arteries that are clogged, stopping the flow is always inconvenient and sometimes dangerous.

Whenever a suspension, composed of discrete particles dispersed in a liquid, flows through a confined geometry, clogging can occur. Channels or constrictions, such as the pores of a filter, can be microscopic, or as in pipes transporting water and in log jams that form under a bridge, they can be macroscopic. As a result, the phenomenon occurs in many environments and scales, as illustrated by the examples in figure 1.

Clogging is problematic in many engineering systems. The blockage of inkjet printer nozzles by colloidal particles impairs their performance. Similarly, the frequent formation of clogs in nozzles used to dispense fiber-filled polymer inks in extrusion-based additive manufacturing processes limits the concentration of fibers that can be used in three-dimensional printing.[1] Another problem—encountered in bioengineering—is the presence of protein aggregates in solutions of monoclonal antibodies. The aggregates threaten the reliability of autoinjection devices that allow patients to self-administer medicine.

On larger scales, clogging is detrimental to water sustainability. Many arid regions rely on underground water as their primary source of fresh water. Those aquifers are either naturally recharged by precipitation or artificially recharged by the redirection of surface water. Both cases depend on water flowing through porous rock and sediment, as in a filter. Over time, cycles of drainage and recharge can cause fine suspended particles to block the aquifer. The progressive clogging increases the energy cost of extraction and recharge and reduces the operational life of the aquifer.

Aquifers are not the only water resource that can be crippled by clogging. Roughly 70% of the water in the US is used for irrigation. And microirrigation, which uses a series of small targeted emitters to water crops, is at least 50% more efficient than sprinkler and furrow irrigation, both of which lose a lot of water to evaporation. Given its superior efficiency, one might expect drip irrigation to be a popular choice. But less than 10% of irrigated land in the US uses drip irrigation, in part because of its susceptibility to clogging, often by suspended sediment, fertilizers, and biofilms of microorganisms, such as algae and bacteria.[2]

Civil engineering is another field in which clogging presents many challenges. As cities have grown, their infrastructure must continually handle more waste. In particular, sewers that channel wastewater and storm runoff allow cities to maintain sanitary living conditions and protect against flooding during periods of high precipitation. The sewers are typically accompanied by inlet and outlet grates that prevent people, animals, and other large objects from entering the sewers. Over time, those grates can become blocked by moss, dirt, leaves, and all kinds of urban trash, such as plastic bags and cardboard.[3] When sewers clog, they cannot handle their designed throughput, which can potentially result in significant flooding. Maintenance of sewer inlets and outlets is thus essential.

Clogging also poses a significant challenge in disease prevention and medicine. A blood clot is an aggregate of platelets and red blood cells, which can block the constriction of a blood vessel and restrict flow.[4] Those clots can become dislodged and clog elsewhere, potentially resulting in a stroke or heart attack, which can be painful, debilitating, and even lethal. Medical devices, such as catheters or bile-duct stents, can likewise be prone to clogging from the buildup of bacterial aggregates, often requiring surgery to remedy. Although they resemble the clogging of nonliving systems, bioclogs are highly complex and largely beyond the scope of this article.

Instead, we focus on clogging by particles in liquid suspensions. But particles in air, where the interstitial phase is negligible, can also clog. Perhaps the most classic case is that of granular flow, such as grains draining through a silo. Grains such as wheat are rough and irregularly shaped. And with no liquid to lubricate them, they interact frictionally with each other, which may easily lead to jamming at the silo's outlet. The effect worsens with high humidity as the grains become more cohesive. When silos clog, the stuck grains must often be manually cleared, either with a long pole, which can be dangerous, or with an air cannon, which is much more expensive.

The clogging of grains shares many similarities with the clogging of active particles, such as cars, people, or livestock passing through a constriction. When sheep enter or leave their corral, they often overload the passage and prevent other sheep from moving. Such overloads are important to consider when planning for evacuation from large buildings and event spaces.

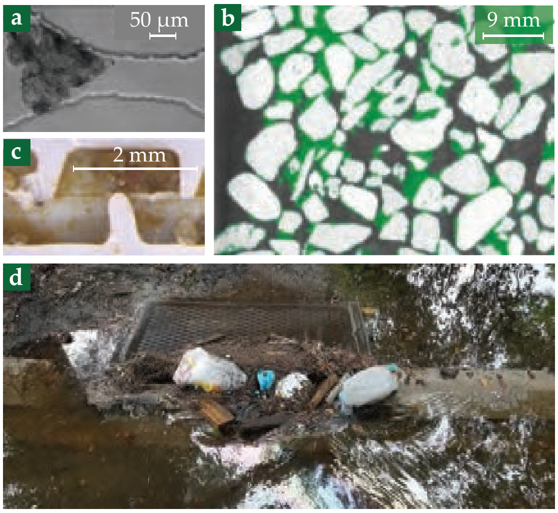


FIGURE 1. CLOGGED SYSTEMS at different scales. **(a)** A protein aggregate is stuck in a microfluidic constriction smaller than the aggregate size. (Adapted from ref. 7.) **(b)** X-ray computed tomography of the internal structure of a porous medium, in which suspended solid particles have flowed for several hours. Trapped particles are shown in green. (Adapted from Y. Tang et al., *Granular Matter* **22**, 37, 2020.) **(c)** An irrigation emitter has become clogged by fine sediment sticking to surfaces along the flow path. (Adapted from ref. 2.) **(d)** A sewer drain outlet is clogged by plastic bags, trash, and other debris. (Courtesy of Bart Everson, CC BY 2.0.)

People tend to rush toward an exit during an emergency, for instance. Emergency exits can become a bottleneck if too many people try to pass through at once.[5] In the best case, the exit limits how quickly people can evacuate. In the worst case, people may try to force each other through or even trample one another. (See the Quick Study by Arianna Bottinelli and Jesse Silverberg, Physics Today, September 2019, page 70.)

The physics that governs how flowing particulate suspensions clog a system has become an increasingly active research topic—partly because of the problem's complexity. It spans many length scales (from bacteria to boulders) and time scales (from less than a second to years), and it often requires sophisticated equipment to study. Yet predicting when clogging is likely to occur can lead to new design principles and improve reliability. One of the first steps in tackling the topic is to categorize its dynamics.

A tale of three mechanisms

Generally, clogging mechanisms include sieving, when particles are too large to pass a constriction; bridging, when particles jam each other at a constriction and form a stable arch; and aggregation, the successive deposition of small cohesive particles at a constriction.[6] One or two of those mechanisms may be more common in certain systems, but generally all of them happen during the clogging process, as can be seen in figure 2, which adds to the complexity of establishing a general picture.

To understand clogging, keep in mind some common parameters. The particles' size, shape, and deformability, and the geometry of the system in which the suspension flows influence the clogging dynamics. When a particle is quasi-spherical, its size is often described by a diameter D, though more complex shapes, such as fibers or particle aggregates, can also jam a constriction. Equally important is a constriction's minimum dimension W, which could represent the diameter of a filter pore, a reduction in a pipe's cross section, or the width of an emergency exit. The volume fraction Φ of particles in a suspension also influences the probability of clogging. (Think about volume fraction as how close, on average, the particles are to each other.) Finally, the physical and chemical properties of the particles—for instance, their roughness, adhesion, and cohesion—also play a crucial role, especially in determining which mechanism dominates.

The combination of those mechanisms establishes a unified framework for the physics of clogging. Nevertheless, each one can be isolated for a more fundamental understanding.

Too big to fit

The simplest case of clogging happens when one dimension of a particle is too large for the particle to pass a constriction. For a rigid spherical particle of diameter D passing through a constriction of width W, clogging occurs for $W/D \leq 1$. That mechanism is one of the most common, as it is routinely used to separate small particles from larger ones with a screen. Because most industrial suspensions are polydisperse—made of different particle sizes—the presence of only a few large particles in the suspensions can greatly hinder the performance of a system with constrictions.

Although sieving seems to be a simple mechanism for spherical particles, it becomes less straightforward when more complex shapes are involved. Anisotropic particles, such as fibers or leaves, may enter a system in a given orientation but later clog the channel after shifting position. That's why gutters must be periodically cleaned. For anisotropic particles, an approach to predict the clogging of a constriction of width W is to characterize the particles' shape using so-called Feret diameters, which entail measuring an object along a specified direction. An anisotropic particle has a maximal Feret diameter L_F along its largest dimension and a minimal Feret diameter l_F along its smallest dimension, as illustrated in figure 3a.

Such a rigid particle will always clog by sieving at a constriction if $l_F > W$, and it will never clog if $W > L_F$. For an intermediate size constriction, $l_F < W < L_F$, however, the particle's orientation determines what happens when the particle reaches the constriction. If the particle passes the constriction with its minor axis aligned with the flow, it won't clog, but if it's aligned perpendicularly to the flow, sieving will occur.

Another complexity arises when the particles are deformable. If larger than the constriction, such a particle can still enter it and may even squeeze through when a large enough pressure is applied,[7] as illustrated in figure 3b. That situation is particularly relevant to biological applications that include deformable vesicles, cells, and protein aggregates.

Clogging by sieving can be described as a random process that depends on the number of particles larger than the constriction or pore sizes. For systems with many pores, sieving can usually be modeled as a Poisson distribution, in which the clogging of each pore represents an independent event. The probability of that event depends on the size distribution of suspended particles.[8] Overall, the most common way to prevent sieving is to stop large particles from entering the system—usually through upstream filtration. Even then, sieving still happens at the filter, where small pores prevent the passage of large particles.

Too crowded to pass

For dilute suspensions—typically of a volume fraction smaller than a few percent—particles tend to pass through constrictions

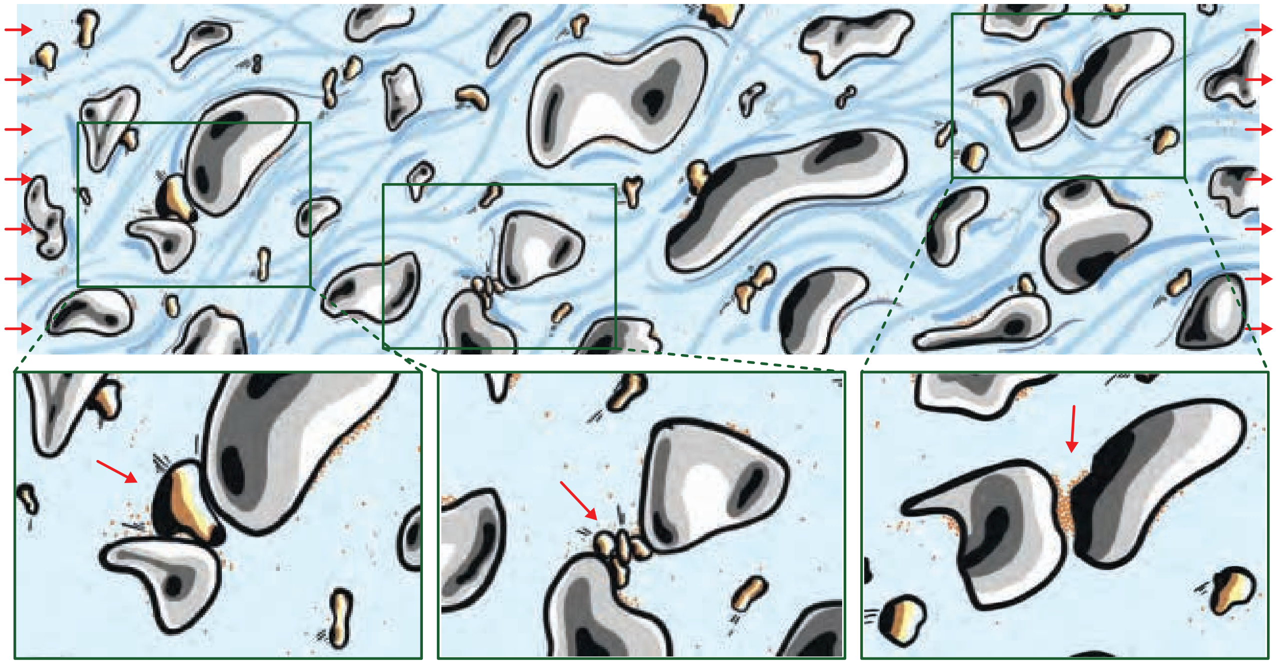


FIGURE 2. A SUSPENSION of particles flowing in a confined system, like this porous medium, can become clogged at constrictions through different mechanisms that depend on the properties of the particles and pores. From left to right: Sieving by a particle with one dimension larger than the width of the constriction; bridging by the formation of an arch of particles arriving simultaneously at the pore; and progressive aggregation on the channel walls of small particles that eventually clog the pore. Red arrows mark the predominant flow direction. (Courtesy of R. S. Sharma.)

one at a time. But clogging by bridging generally occurs at higher volume fractions when particles arrive at a constriction simultaneously.[9] If too many particles arrive at once, they can jam each other, spontaneously forming a bridge, as illustrated in the three time-lapsed frames in figures 4a–c. An arch is formed when constrictions are of size *W*—larger than the particle size *D*, but of the same order of magnitude.

For spherical noncohesive particles, researchers have observed bridging experimentally only when $W/D < 5$. For highly anisotropic particles, such as fibers, flakes, or cohesive particles, such as powders, that threshold can be higher because of differences in the particles' arrangement and the resulting force network at the bridge. Clogging by bridging shares many features with the clogging of silos, in which particles smaller than the constriction can form stable arches. The main difference is that silo clogs occur only at large packing densities because the dry particles are subject to gravity.

An approach commonly used to describe such systems is to measure the number of particles *s* or the volume of suspension that flows through the constriction before a bridge is formed. The two relevant quantities are the average number of particles $\langle s \rangle$ (or the average volume of suspension) before clogging occurs and the resulting distribution of the number of particles escaping. The latter exhibits an exponential decay because of the constant probability of clogging during the flow.[10] As a result, the probability of clogging also follows an exponential distribution, illustrated in figure 4d.

An approach to predict the average number of particles, or the average volume of fluid, flowing through the constriction before bridging occurs is to assume a random distribution of particles in the channel and imagine that a sufficient number of them reach the constriction at the same time. For instance, in a 2D system and particles of diameter *D*, a constriction of width *W* will clog by bridging if the number of particles arriving simultaneously is greater than or equal to floor($W/D + 1$). For a given particle density, it is then possible to show that before the system clogs the average volume of fluid will increase if the width of the constriction is increased (figure 4e). Similarly, clogging will be delayed if one decreases the volume fraction of the suspension.[11] Clogging by bridging exhibits a particular feature of intermittency, in which the flow clogs and unclogs periodically, when some perturbations exist in the system. Those conditions provide a framework to minimize clogging by bridging.

Too sticky to flow

Very small particles can stick to surfaces and to each other. If a particle attaches to the wall of a constriction, that constriction becomes a little smaller,[12] as shown in figure 5. When the deposition happens repeatedly, particles aggregate, eventually leading to severely reduced constrictions that are more susceptible to both bridging and sieving.[13] Although initial deposition depends on the interaction between the particle and the wall, interparticle interactions can significantly increase the aggregation. The particle–fluid interaction is also an important parameter, because the flow determines how quickly particles will arrive and the shear forces they experience. And that force may influence whether or not the particles can be eroded.

For particles such as sediment or microplastics, aggregation is primarily caused by van der Waals forces, which are short-range forces from distance-dependent interactions between atoms and molecules. Other small particles, such as bacteria and algae, can also be sticky. While still subject to van der Waals forces, bioparticles are often decorated by a variety of adhesion molecules on their exterior, which allow them to attach to surfaces

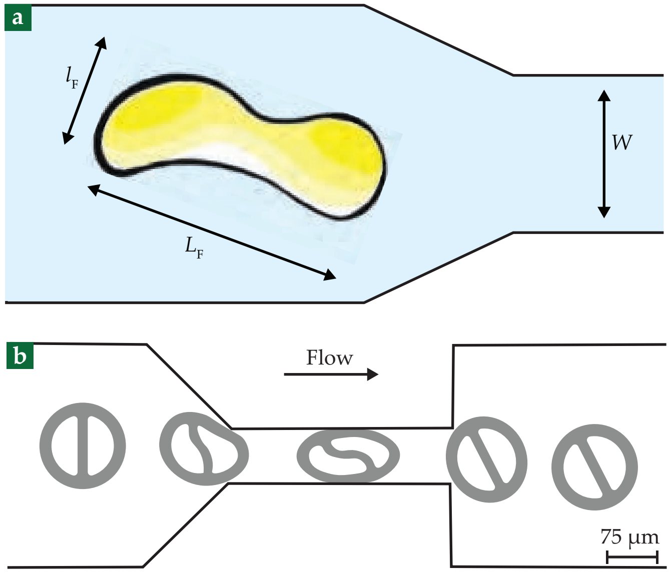


FIGURE 3. GETTING A GRIP on deformable particles. **(a)** Anisotropic and deformable particles have minimal and maximal Feret diameters l_F and L_F, respectively—a Feret diameter refers to the measure of an object along a specified direction. The particle shown here flows toward a constriction of width *W*. (Courtesy of R. S. Sharma.) **(b)** A flexible particle can squeeze through a constriction smaller than its size if the input pressure is large enough. (Adapted from L. Chen, K. X. Wang, P. S. Doyle, *Soft Matter* **13**, 1920, 2017.)

or each other with greater force than abiotic particles. Once attached, many bioparticles can also multiply along the surface while excreting chemicals that serve as a glue to bind them there.

Sieving and bridging are governed primarily by interactions caused by spatial arrangements, but aggregation is more complicated and sensitive to chemistry. For particles in pure water, aggregation is often limited by an electric double-layer repulsion. But pure water is extremely rare outside of specialized applications, such as wafer processing and scientific research. Changes in pH or the addition of salt increase ion availability in suspensions and can drastically reduce double-layer repulsion, allowing van der Waals interactions to dominate. Even small concentrations (<0.1 mol/l) of monovalent salt can significantly increase the aggregation rate.[14] Usually, clogging by aggregation can be observed even for constrictions much larger than the particle diameter and for very dilute suspensions. This clogging mechanism, however, usually takes place over long periods of time—from hours to years.

The flow must go on

Because clogging is such a problem in a broad range of fluid systems, much of contemporary research investigates new techniques for preventing or mitigating it. Just as different systems experience different clogging dynamics, techniques to prevent sieving, bridging, and aggregation also vary depending on the predominant mechanism at work.

A common way to prevent sieving is to filter particles upstream. Most suspensions, however, are not ideal. Many particles, such as cells or protein aggregates, are anisotropic in shape and may be deformable. Nonetheless, those imperfections are actually opportunities to prevent clogging.

For deformable particles, or rigid particles flowing through a deformable constriction, elevated pressure can squeeze them through the passageway.[7] When that happens, the deformation of the particle, constriction, or both results in an adjustment of *W*/*D* that allows the particle to pass. Humans and other mammals rely on that phenomenon, as red blood cells must deform to flow through capillaries. In fact, numerous diseases alter the deformability of red blood cells, increasing the risk of blood clots and stroke if those cells become so stiff that they can no longer flow freely through small capillaries.

Highly anisotropic particles may clog a constriction in one orientation but flow freely in another orientation. To prevent sieving in that case, the particle must become aligned so that it can pass the constriction. For some shapes, particularly those with high aspect ratios, hydrodynamics provides a mechanism. At low volume fractions, slender particles subject to a shear flow tend to align their longest dimension with the flow direction. At higher volume fractions or turbulent, high flow rates, however, the motion of anisotropic particles in confined systems is less certain and remains an active research topic.

The simplest way to prevent bridging is to ensure that the suspension volume fraction is sufficiently low and the constriction is sufficiently large. In the case of grain silos or emergency exits, however, achieving either may be impractical or even impossible. Fortunately, other ways help prevent suspended particles from bridging. Indeed, once a bridge has formed in a static system, all the kinetic energy dissipates and the bridge remains stable with an unclogging probability of zero. But the stability of that bridge can be broken by introducing perturbations to the system.

One way to periodically perturb dry granular media is through vibrations using a piezoelectric device or, in the case of a liquid suspension, by adding fluctuations to the flow. Researchers have shown that, depending on its frequency and intensity, such vibration increases the unclogging probability. Thus, systems that rely on vibration to break bridges often exhibit an intermittent flow, which reflects both clogged and unclogged states.

Another passive method to prevent bridging in silos is to place an obstacle just upstream of the constriction. Researchers have shown that proper placement reduces the bridging probability by a factor of 100 without reducing the flow rate. Furthermore, that technique works for both passive particles, such as grains in a silo, as well as active particles, such as sheep or humans rushing through a constriction.[15] The addition of an obstacle reduces the pressure upstream of the outlet and modifies the conditions for bridging, with the ultimate effect of reduced bridging probability, even for dense suspensions.

Preventing aggregation has proven to be a more difficult task thus far. Nevertheless, there are some promising techniques. For a long time, the best way to prevent aggregation has been to introduce chemicals into a system. Some examples include adjusting the system's pH to reduce mineral precipitation in hard water, or adding biocides, such as chlorine, which kill bacteria and algae and prevent them from proliferating. Adding chemicals is far from ideal, however, as many chemicals are only compatible with specific systems. Furthermore, adding chemicals to large flows presents an environmental risk,

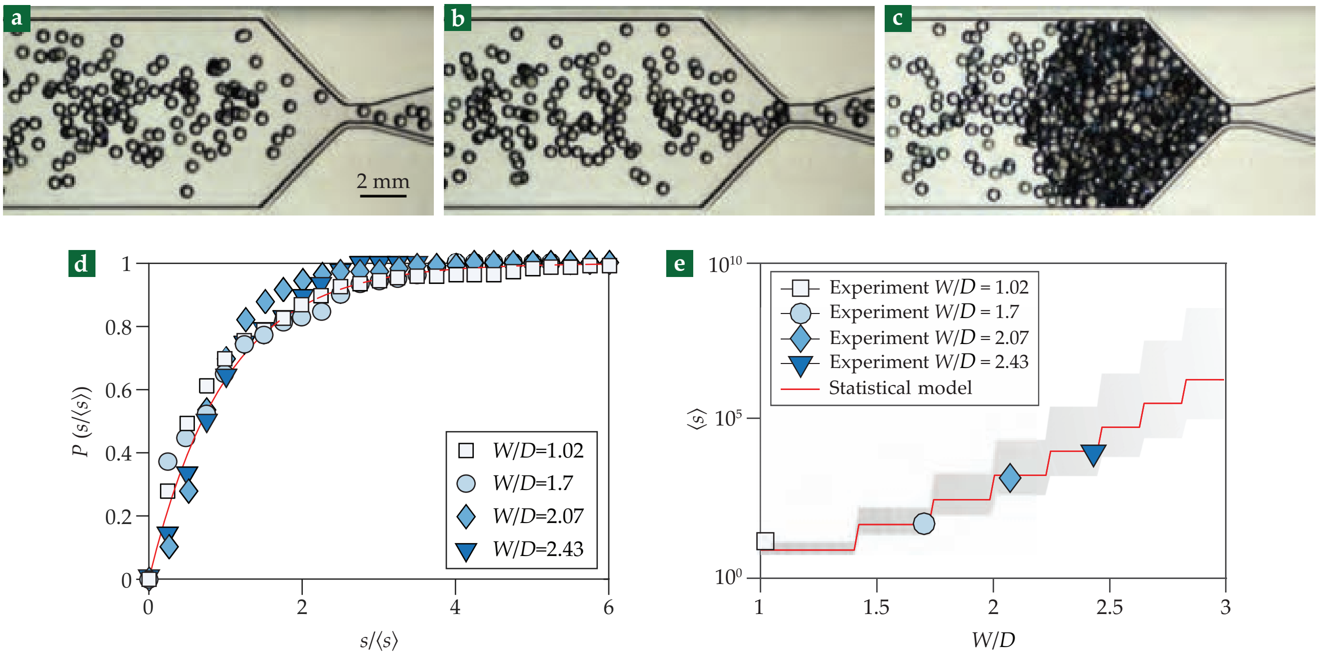


FIGURE 4. PARTICLES BRIDGE at a constriction. **(a)** Particles flow through the channel until **(b)** a sufficient number of them reach the constriction at the same time to clog it. **(c)** The clog prevents the flow of particles, which then form a filter cake. **(d)** This chart shows the probability *P* of clogging after *s* particles have escaped the constriction for differently sized particles *D* and constriction widths *W*. **(e)** The average number of particles <*s*> that can pass through the constriction before clogging it varies with the size ratio of the constriction width to particle diameter, *W*/*D*. (Adapted from ref. 10.)

because they could run off and invade nearby ecosystems. More robust techniques are therefore being investigated.

In some cases, surface treatment can help prevent aggregation. Modifying the surface roughness or surface energy—and often both—significantly reduces the aggregation rate. But it is difficult to develop surface treatments that maintain anticlogging properties for a long time, because of either erosion or fouling of the surface.

Another technique that researchers have proposed leverages capillary forces to remove particles.[16] For instance, passing a slow-moving bubble through a system can do the trick. As the bubble passes over particles, it exerts an inward capillary pressure that can peel particles from the surface. The process of capillary peeling has been shown to work for both inert particles and bioparticles. And although researchers have demonstrated the process for short-term removal of particles, no one has yet tested it as a long-term solution to curb aggregation. Additional methods to locally increase the fluid shear are being developed for future applications.

Finally, yet another promising solution to mitigate clogging is to incorporate pulsating flows into the system.[17] Studies dating as far back as the 1980s report on their anticlogging potential. Over time, researchers have observed that the unsteady shear environment associated with pulsatile flows may help mitigate all three mechanisms of clogging. With pulsation, anisotropic particles can rotate so that they pass a constriction and do not sieve. And although particle bridges are often stable in a steady flow, pulsation can reorient or break apart a bridge and restore flow to the channel. The mechanism is similar to how vibrations prevent bridging. Finally, pulsatile flows may mitigate aggregation by periodically increasing the fluid velocity and eroding particles with a temporarily elevated shear.

Eyes on the future

Many different systems can clog in distinctly different ways, so no one-size-fits-all solution exists, and a lot of work remains to obtain a unified picture of how it happens. Understanding the impact of different parameters on clogging is important for designing resilient systems, thus saving time, money, and energy.

Ideally, as our understanding of clogging becomes increasingly comprehensive, general guidelines for its risk can be established. For bridging and sieving, we know that the size ratio W/D and the volume fraction Φ are the dominant parameters

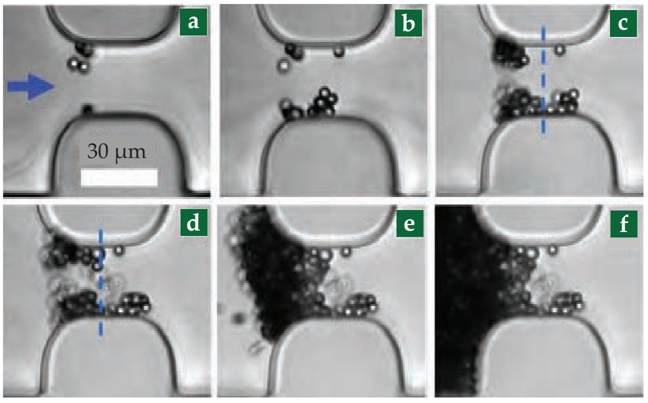


FIGURE 5. THE SUCCESSIVE DEPOSITION of particles at a constriction eventually clogs it. Images correspond to the following times: 4, 9, 12, 16, 22.5, and 24 minutes. Every minute 68 000 particles pass through the pore. The blue arrow signifies the flow direction and dashed blue lines correspond to the middle of the pore. (Adapted from ref. 12.)

that determine whether or not a system will clog. Therefore, it may soon be possible to describe a phase diagram for the phenomenon. Eventually, parameters such as roughness, particle anisotropy, and deformability may be incorporated into such a diagram. The influence of channel geometry, such as the angle of a constriction, needs to be considered as well.

Aggregation is more complex, but it may be handled in a similar way. We can assume that $W/D \ll 1$, which makes the constriction width a less important consideration for an aggregation phase diagram, though it would change the time it takes the channel to become fully clogged. Similarly, increasing the volume fraction Φ should simply increase the aggregation rate. The proper phase diagram would indicate when to expect aggregation, given the competition between attachment forces—due to a combination of van der Waals forces and an additional adhesive force—and erosion forces from the shear exerted by the fluid. If the erosion force exceeds the attachment force, we would expect minimal aggregation or an upper limit to the size of the average aggregate. But if the attachment force exceeds the erosion force, we would expect continual aggregation and eventually a complete clog.

To fully understand those processes requires an interdisciplinary approach. Clogging can be sensitive to hydrodynamics, biology, chemistry, and physics all at once. Suspensions may contain billions or more of polydisperse particles with varying properties, and their specific locations in the suspension are unknown. Thus, as in turbulence modeling, advanced clogging models require extensive validation and iteration to be predictive. And although simplified systems are essential for isolating specific aspects of clogging, complex systems must be investigated as well. Real suspensions are polydisperse and flow through a wide variety of geometries. Field research and case studies that focus on such real systems should be extremely useful for helping scientists design relevant experiments while also serving as validation sources for clogging models.

REFERENCES

1. B. P. Croom et al., *Addit. Manuf.* **37**, 101701 (2021).
2. S. Bounoua et al., *Irrig. Sci.* **34**, 327 (2016).
3. J. P. Leitão et al., *Stoch. Environ. Res. Risk Assess.* **31**, 1907 (2017).
4. A. L. Fogelson, K. B. Neeves, *Annu. Rev. Fluid Mech.* **47**, 377 (2015).
5. N. Shiwakoti, X. Shi, Z. Ye, *Saf. Sci.* **113**, 54 (2019).
6. E. Dressaire, A. Sauret, *Soft Matter* **13**, 37 (2017).
7. C. Duchêne et al., *Soft Matter* **16**, 921 (2020).
8. A. Sauret et al., *Appl. Phys. Lett.* **105**, 74101 (2014).
9. I. Zuriguel et al., *Sci. Rep.* **4**, 7324 (2014).
10. A. Marin et al., *Phys. Rev. E* **97**, 021102(R) (2018).
11. N. Vani, S. Escudier, A. Sauret, *Soft Matter* **18**, 6987 (2022).
12. B. Dersoir et al., *Microfluid. Nanofluid.* **19**, 953 (2015).
13. R. Jäger, M. Mendoza, H. J. Herrmann, *Phys. Rev. Lett.* **119**, 124501 (2017).
14. N. Delouche, A. B. Schofield, H. Tabuteau, *Soft Matter* **16**, 9899 (2020).
15. I. Zuriguel et al., *Phys. Rev. E* **94**, 032302 (2016).
16. S. Khodaparast et al., *Environ. Sci. Technol.* **51**, 1340 (2017).
17. B. Dincau et al., *Soft Matter* **18**, 1767 (2022).

PT

The Petronas Twin Towers form the centerpiece of the famed skyline of Kuala Lumpur, Malaysia's capital. (Photo by Sean Pavone/Alamy Stock Photo.)

Clarissa Lee is a junior fellow at the RWTH Aachen University's Käte Hamburger Kolleg: Cultures of Research in Aachen, Germany. Her work focuses on the philosophy, sociology, and history of science, technology, and media in Southeast Asia.

MALAYSIAN PHYSICS AND THE MAKER ETHOS

Clarissa Ai Ling Lee

Far removed from traditional research centers, Malaysian physicists forged a scientific culture that prioritized innovation and reuse.

Although the first physics department in present-day Malaysia was founded only in 1961 at the University of Malaya (UM), the country began its foray into modern physics in the mid 1950s, at the same time that it was beginning to form a national identity. That was a complicated process: The Federation of Malaya achieved independence from the UK in 1957 and united with three other former UK colonies—Singapore, Sabah, and Sarawak—to form Malaysia in 1963. (Singapore agreed to a separation in 1965 and is now an independent city-state.)

UM, the country's first university, was founded in 1949 when colonial officials merged the King Edward VII College of Medicine with Raffles College. With the university, they hoped to provide a common Malayan identity that would represent a compromise between the colony's different ethnic groups.[1] That compromise began to take shape after the 1942–45 Japanese occupation during World War II, when the UK returned to a politically changed territory that was suddenly less subservient to its colonial policies. Surviving records of student activities during the eight years before Malaya's independence in 1957 illustrate increasing nationalist and political agitation on UM's campus.

For UK administrators, the university was to be the diplomatic link between the new administration of an independent Malaya and the UK. Science was seen as a neutral ground devoid of political undercurrents: Malayan scientists seem to have rarely participated in anticolonial protests, nor did they appear to engage in other forms of political activism. The apolitical nature of the Malayan scientific enterprise contributed to a culture that increasingly valued empirical research over work that was more philosophical. Academic science departments were set up to both carry out bread-and-butter goals and deal with nation building in "administrative, legal, and economic systems."[2]

The lack of a clear connection between philosophy and science shaped a more pragmatic attitude toward physics. But that did not translate into the straightforward approach of "shut up and calculate" because the limited resources available in Malaysia meant that physicists needed to find ways around knowledge black boxes. That need gave birth to the "maker ethos," a term scholars have coined to describe do-it-yourself cultures and practices of crafting, fabricating, and repurposing technologies for everyday applications. In maker cultures, people of diverse backgrounds come together to nurture hands-on solutions and bring social values into the practice and circulation of knowledge. In so doing, they offer a counterargument to the widely held idea that technology alone determines the path of history.[3]

The maker-ethos framework enables scholars to understand how different and seemingly antithetical practices from varying scientific knowledge traditions—modern Western physics and medieval astronomy in the case of Malaysia—can coexist. It explains how Malaysian physicists connected local needs with the international community and how the government's call for "technological action" became the impetus to encourage the pursuit of medieval astronomy, which was deemed suitable to fulfill a religious ritual important to the country's Malay Muslim community: the observation of lunar phases to properly determine the start and end of the fasting month, Ramadan.

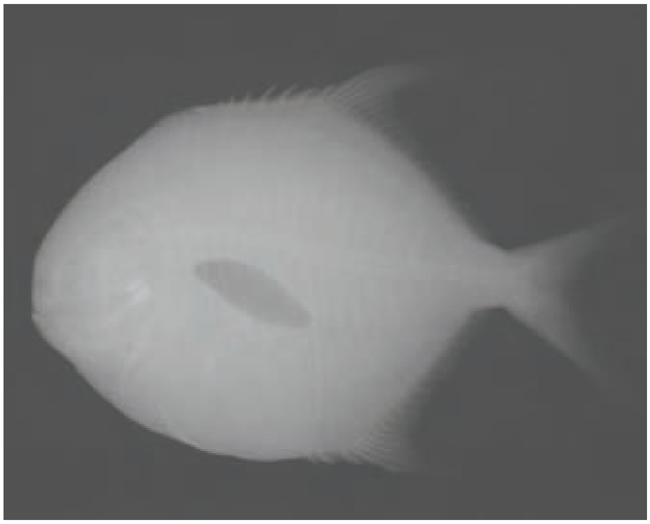


FIGURE 1. A REPRODUCTION of the first x-ray photograph made in present-day Malaysia. The image, taken in 1897, is of a pomfret fish. (Courtesy of Mohd Hafiz Mohd Zin and the Science University of Malaysia.)

Physics from the ground up

Beginning in the late 1920s, well before the first physics department was established at UM, selected schools in pivotal centers of the Malayan colony started offering rudimentary physics courses as part of basic science education. Aside from that, physics was taught on a need-to-know basis. Starting in the late 19th century, for example, doctors and clinicians began receiving training in basic physics principles, which were necessary for handling medical x rays and radioisotopes[4] (see figure 1).

The establishment of UM did not happen in a vacuum: It was one of several research institutions set up by colonial authorities that aimed to reap economic benefit from basic science. Before modern Malaysia was formed with its current borders in 1965, Malayan physicists primarily collaborated with their former PhD advisers in the UK and Australia or participated in international collaborations, such as the International Geophysical Year of 1957–58 with UK assistance.[5] Only in the 1970s did Malaysian physicists begin to spend time doing research abroad.

By that time, Big Science was taking off in more resource-capable countries. Like other postcolonial nations, Malaysia focused on tabletop experimental physics, which remained crucial to facilitating both applied and exploratory research. As it grappled with how to fund research on a limited budget, the country began to face the famous tension between basic and applied science.[6] At first the divide was not so pronounced, but by the 1980s, budget cuts at universities led to increased pressure to develop research that could support itself through private and industrial funding.

Physics and industry

By 1965, newly trained Malaysian physicists who had received their PhDs abroad were starting to return home. One of those individuals was Thong Saw Pak, who spent time in China in the late 1950s as a postdoctoral researcher before becoming the first chair of the department of physics at UM's campus in Kuala Lumpur. (Before Singapore's secession in 1965, the university had two campuses: one in the present-day city-state and one in Kuala Lumpur. The former is now the National University of Singapore; the latter is modern-day UM.) In a 1962 lecture, Thong argued that physics contributed to technological advancements and improvements in quality of life. For him, physics was central to industrialization. Similar lines of thought were in a 1966 speech given at UM by Karl Emeléus, a visiting physicist from Queen's University Belfast.[7]

Both speeches illustrate how the bridge between basic science and technology stemmed from a philosophy that prioritized inventiveness as a way to best achieve development goals. That made poorer, postcolonial states dependent on wealthier states, because richer countries sought to cultivate knowledge producers who would create proprietary science and technology that would benefit foreign investments. It also implicitly allowed wealthy states to dictate which research programs were valued and prioritized.[8]

Making physics

In the 1960s plasma physics was seen as the bridge between theoretical quantum physics and the needs of industry. And plasma physics required equipment that could be used in other, related areas of physics that were also in vogue at the time, such as nuclear, radiation, and ionospheric physics. For that reason, Malaysia began to develop expertise in particle, laser, and condensed-matter physics in the 1970s and 1980s. Because the country had few physicists and only three universities with a physics department—UM, the National University of Malaysia (UKM), and the Science University of Malaysia (USM)—it was unsurprising that scientists prioritized research programs that straddled multiple fields and, in the case of experimental and applied physics, allowed them to reuse available equipment (see figure 2).

In terms of the number of trained Malaysian experts available, plasma physics, nuclear physics, and—by the 1980s—laser physics dominated. With the assistance of the United Nations University in Tokyo, which facilitated research training, experimental physicists from UM were at the forefront of fostering collaboration between developing countries by training physicists from Egypt, India, Indonesia, Nigeria, Pakistan, and Sierra Leone. The Malaysian scientists taught them how to plan and design methods for developing their own experiments that would contribute to their national research programs. Doing so involved balancing costs, matching instruments with each country's research and teaching programs, and considering how equipment might be adapted for research in related subdisciplines.

After receiving startup funding, the physicists were responsible for deploying their newly gained technical knowledge in their home countries (see the article by Sing Lee and Chiow San Wong, Physics Today, May 2006, page 31). They were encouraged to seek locally accessible solutions, be inventive in their research planning, and craft modular setups that could be repurposed for different experiments.

Between 1949 and 1960, physicists at UM's campus in Singapore pursued tabletop research in ionospheric and nuclear physics. Making use of instruments and data from the colonial period along with some equipment they were able to purchase, they focused largely on measurement and calibration studies. Some of their publications during that period demonstrated how cosmic-ray studies in atmospheric physics, for example, could be readily transferred to plasma physics.

Because most of them had not trained with physicists with established research programs in then-new subfields of physics, such as relativity theory and quantum mechanics, the first generation of Malaysian physicists was largely unfamiliar with those topics. For that reason, the research they conducted typically focused on reproducible research that was likely to be pedagogically and industrially useful. Those physicists also had to ensure that they were sufficiently trained with instruments so that they could help set up the laboratories for the new universities established from the 1960s onward, such as UKM and USM.

The maker ethos helps scholars understand all the above phenomena, especially Malaysian physicists' backgrounds, their choice of research programs, and the international collaborations they established. Although the networks of scientific exchange that were set up in the colonial and early postcolonial periods had involved other physicists from the Commonwealth of Nations, by the 1970s they expanded to include physicists in other developing Southeast Asian countries, such as Indonesia and the Philippines, and those in newly industrialized East Asian countries, such as Japan.

Published proceedings from regional physics conferences from the 1970s onward provide a glimpse into the tension between science that prioritized matters of domestic importance and science seen as having more international panache. They show that what counted as physics was not parceled out in neat terms—the maker ethos that underpinned the development of scientific knowledge meant that the type of scientific orthodoxy that privileges only certain topics as a legitimate focus for physics research did not exist in the postcolonial Southeast Asian context.

Experiment or theory?

Under the maker ethos, the distinction between theory and experiment is blurred. Malaysian plasma physicists of the 1970s saw themselves as contributing to research on nuclear fusion through their experiments. Malaysian experimentalists did work with local theorists—at the time, USM already had a small program in theoretical plasma physics—to ensure that their experimental designs were rigorous. But their mathematical scaffolding was still limited and largely used to mediate between an analytical conceptual framework and numerical methods.

FIGURE 2. SCARCE RESOURCES often forced Malaysian physicists to reuse or repurpose old laboratory equipment. This image depicts a closet full of old apparatus at the Science University of Malaysia that dates from the university's early years.

Not all theoretical work was done to aid local experimentalists, however. Starting in the late 1970s, a small group of Malaysian theorists who worked mainly on particle physics began to consider smaller research questions foundational to problems in that field. Their work helped clean up some of the paradoxes in larger physics models.

Through interviews, I have learned that physicists of that generation (see figure 3) tended to work in isolation and often got stuck on problems that were no longer viewed as trendy in prestigious centers of theoretical physics. One could interpret that obsession as Malaysian physicists' attempting to focus on improving existing theoretical models and revealing their inherent flaws. But the technical language they used in their papers would have been incomprehensible to the first generation of Malaysian physicists, who weren't trained in the required mathematics. Ironically, one would often find such articles published in conference proceedings alongside other papers that had a more pragmatic bent and were lighter on mathematics. Considering that those conferences were attended

FIGURE 3. A 1974 GROUP PHOTO of the University of Malaya physics department, including professors, staff, and students. (Courtesy of Sing Lee.)

mainly by physicists working on more pragmatic topics, one may wonder as to whom the intended audience of the theoretical articles was.

Until the 21st century, Malaysian universities didn't have any experimental research groups in particle physics, so, as I've learned in interviews, experimentalists either joined collaborations in other countries, including West (and, later, reunified) Germany and Japan, or took on other projects where they could deploy their skill sets. Funds for traveling abroad were minimal, so Malaysian physicists began serving as experts in technical-assistance programs or applied for scholarships from such funding sources as Colombo Plan scholarships and the UK's Commonwealth Scholarships.

Unlike the first generation of Malaysian physicists, who were able to carry on the research programs they had trained in despite resource limitations, younger physicists trained abroad found it increasingly difficult by the late 1980s to continue their research upon their return home. That meant that physicists who trained elsewhere often chose not to return to Malaysia to avoid disrupting their career; the resulting brain drain contributed to discontinuity in the generational flow of physics research.

Although Malaysia had no formalized national policy regarding science and technology before 1985, its annual national development plans included provisions to fund recruitment and training of Malaysian scientists and to build scientific infrastructure. The physicists saw their research as an opportunity for economic and social advancement: Several of the third-generation physicists trained in the 1970s and 1980s came from working-class backgrounds. Most of them were men, but some were women, including the theoretical physicist Rosy Teh (who spent six months as a postdoc at CERN). Because physics was viewed as a service to the nation, domestic accomplishments were seen as more worthwhile than recognition abroad. That tension can be seen in the arguments made by physicists themselves, who believed that Malaysian physics should be more focused on economic development than driven by curiosity.[9]

Scarce resources and limited infrastructure required Malaysian technical experts to inventively adapt or reconstruct equipment,[10] which meant that scientific research involved crafting, hacking, and repurposing (see figure 4). But leading members of the community resisted embracing more pragmatic short- to mid-term utilitarian goals for physics and instead placed importance on practicing physics as it was done in elite centers.[11] That tension stemmed from the way physics had been embraced in postcolonial Malaysia and its irreconcilability with the heterogeneous knowledge cultures present in the country's varied and diverse ethnic groups. The other tension came from following a policy of modernization that was alienated from everyday struggles.

Nevertheless, the sciences, including physics, were relatively meritocratic in their recruitment and retention of talent during that time. Although Malaysian physics was at first dominated by ethnic Chinese, that began to change with the founding of UKM in 1970. That institution was set up to overcome the shortage of ethnic Malay technical experts through recruitment and training, and its establishment began the gradual shift of the language of science instruction from English to Malay. Aside from language courses, Malay is now the sole language of instruction at UKM, although English is being gradually reintroduced to appeal to international students.

Given the lack of local scientists who were able to teach science in Malay at a high level, Indonesian scientists were brought in to fill the gap well into the 1970s. The Malaysian government also recruited and sent bright Malay students for training abroad under its New Economic Policy, which led to an increase in the number of ethnic Malay physicists by the

FIGURE 4. A DIFFUSION CLOUD CHAMBER, at left, homebuilt by Andrew J. C. Chong, a Malaysian physicist. At right is a close-up of the object in the chamber: a small piece of the mineral thorianite. Under the influence of the chamber's magnetic field, the thorianite emits alpha particles (thick, straight, short lines) and beta particles (thin, curvy, faint lines). (Courtesy of Andrew J. C. Chong.)

mid 1970s. Later in the century, merit-based affirmative action gave way to increasingly racialized policies in the civil service that privileged Malays over other groups and contributed over time to the Malayization of the public university system and a decrease in scientists from other ethnic groups.

Anachronistic knowledge

One might assume that there would be no connection between indigenous Malayan and Malaysian knowledge traditions and the modern physics introduced into the country in the 20th century. But some in Malaysia argue that the practices of scientific knowledge under the auspices of Islamic beliefs are concomitant with modern Western physics. One such example is *ilmu falak*, or Islamicate astronomy and astrological divination, which are touted by their advocates as being representative of the scientific traditions of the Malay world. In Malaysia, astrological divination is still taught in the *sekolah pondok*, a traditional type of Malay school roughly equivalent to the US one-room schoolhouse. In such schools, students from varying age groups are taught recitation-style by religious teachers.[12] Those schools predate the colonial era and still exist today. They serve mainly as a supplement to the secular school system.

Although the theoretical and practical knowledge—including a heliocentric view of the universe—that those schools taught would be considered legitimate from the perspective of modern physics, the material drew heavily on premodern and early-modern astronomical knowledge.[13] That would not change in the 1980s, when modern astronomy began developing as a field in Malaysia. Religious teachers taught not only theories but also methods of constructing handheld instruments, such as astrolabes that used the lunar calendar to determine dates and times for the performance of religious rituals (see figure 5). I attended a workshop on building a personal astrolabe in 2015 aimed at graduate students studying philosophy of science in the Islamicate world. Although the workshop was conducted by two Dutch scholars, the methods for building and calibrating the astrolabes used the same concepts and source material as those taught at *pondok* schools.

That type of learning can be understood through the maker-ethos paradigm because even though religious teachers at *pondoks* had limited training in mathematics and the sciences, they made a concerted effort to render accessible to nontechnical experts the calculations and measurement techniques that could be used to calibrate calendrical events. Most of that knowledge was derived from religious texts of varying levels of technical detail.

Because many medieval astronomy texts were written in Arabic, they were orally translated by teachers who could read and understand the language. The translation was faithfully written down by their students, who then compiled the notes into collections. Teachers with requisite Arabic knowledge were spread across colonial Malaya, so it was not unusual for more-senior instructors to correspond with younger colleagues on problems requiring technical resolution. In other words, those religious teachers and their pupils hacked scientific knowledge to create useable technologies that enabled themselves to fulfill their religious obligations.

Moreover, that knowledge was not indigenous to the Malay community but was localized to meet its needs. Although traditional astronomical knowledge is no longer taught in religious schools, Islamic-studies departments across Malaysia still teach it. Even if their university has a physics department, Islamic-studies scholars often don't interact with secular physics faculty, although some Muslim physicists occasionally get involved with traditional astronomical work because they are interested in it and therefore contribute their technical knowledge to the collaborative endeavor.

Compared with the more transnational way in which Islamicate astronomy knowledge was transmitted in the mid 20th century, much present-day Malaysian research in that area has become isolated from studies of Islamicate sciences elsewhere. Nevertheless, the fact that Islamicate astronomy, divination, and astrological studies coevolved with secular physics

FIGURE 5. A MODERN-DAY ASTROLABE constructed by Jacopo Koushan in Iran in 2013. Similar instruments are still constructed and used by members of the Malay Muslim community to assist in such religious rites as the determination of prayer times and the direction of Mecca. (Photo courtesy of Masoud Safarniya, CC BY-SA 3.0.)

departments in the 1950s and 1960s despite appearing to have no relationship to each other is representative of the plural forms of modernity in existence.

Modernism and physics knowledge

Early-20th-century physics often served imperialist aims: Practitioners who were trained in university laboratories in imperial metropoles such as the UK, France, and the Netherlands would subsequently serve the needs of colonial administrators by finding permanent appointments in various institutions.[14] Colonial officials encouraged the development of astronomy because methods used to stargaze and canvass terrestrial sites could also be used to survey uncharted lands. That was also why imperial officials encouraged the development of meteorology and geology in the colonies and fostered the establishment, for example, of local weather bureaus. The fields of science that were allowed to develop in the colonies were not just following global knowledge trends but often steered by external political interests. For instance, even though the UK established several sites for astronomy research in colonial Malaya, modern astronomy and astrophysics were never included among key physics subfields in Malaysia until UKM inaugurated the beginnings of a space-studies center in 1990. A national planetarium[15] wasn't established until 1994.

Malaysian physicists did not partake in the early-20th-century philosophical debates that forged the idea that classical and quantum physics were separate knowledge systems. As far as they were concerned, all physics was modern.[16] During the 1950s and 1960s, experiments carried out by local physicists were simple. They focused on replicating measurements, collecting data based on others' experimental designs, and designing experiments that would enable them to test newly acquired or newly built apparatus. The knowledge gained from those investigations was then transferred to local technological and industrial needs. By the 1970s, physicists began supplementing their studies with basic-science experiments in such subfields as nuclear and condensed-matter physics.

Modern physics can be considered to be part of the postimperial postwar order, which aimed to ensure the continued dominance of states with global influence, whose arsenal of scientific knowledge would be bartered with emerging states in the guise of economic development and industrialization. The astronomy of the Malay Muslim community stood apart from that phenomenon. But even though they had different goals, both Islamicate astronomy among the Malay Muslim community and modern physics embodied the maker ethos. In both instances, the maker ethos was crucial to the intellectual confidence of Malaysian physicists, who were beginning to form professional and cultural identities. It allowed them to carve a path for themselves despite their complicated political situation. That same spirit would influence the development of Malaysian physics from the 1980s onward.

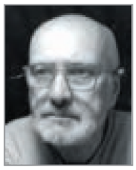
REFERENCES

1. A. M. Carr-Saunders, *New Universities Overseas*, George Allen & Unwin (1961), p. 57.
2. Commission on University Education in Malaya, *Report of the Commission on University Education in Malaya*, His Majesty's Stationery Office (1948), p. 10.
3. S. Nascimento, A. Pólvora, *Sci. Eng. Ethics* **24**, 927 (2018).
4. M. R. bin Hashim, in *Centennial of the X-Ray: An Account of Developments in Radiological Physics and Radiology in Malaya and Malaysia*, A. A. Tajuddin, D. A. Bradley, eds., Malaysian Institute of Physics (1995), p. 1.
5. "Colony's role in world's biggest research project; all eyes (scientific) on skies this year," *Straits Times*, 14 February 1957, p. 6.
6. G. Gooday, *Isis* **103**, 546 (2012).
7. K. G. Emeléus, *The Role of Physics in Industrial Development and Other Lectures*, U. Malaya (1966).
8. M. A. Ynalvez, W. M. Shrum, in *International Encyclopedia of the Social and Behavioral Sciences*, 2nd ed., vol. 21, J. D. Wright, ed., Elsevier (2015), p. 150.
9. C. Singh, in *Peranan fizik dalam negara membangun/The Role of Physics in Developing Countries*, National U. Malaysia Department of Physics (1979), p. 25.
10. K. H. Ong, in ref. 9, p. xi.
11. S. C. Lim, in ref. 9, p. 18.
12. A. A. Mat Zin, in *Sejarah astronomi Islam di Malaysia* (History of Islamicate astronomy in Malaysia), A. A. Mat Zin, ed., U. Malaya Press (2016), p. 17.
13. N. Ahmad Zaki, in ref. 12, p. 27.
14. L. Pyenson, *Empire of Reason: Exact Sciences in Indonesia, 1840–1940*, E. J. Brill (1989).
15. Z. Z. Abidin, M. F. Asillam, J. Y. Koay, *Nat. Astron.* **4**, 1115 (2020).
16. R. Staley, *Isis* **96**, 530 (2005). PT

Accelerating astrophysics with the SpaceX Starship

NURPHOTO SRL/ALAMY STOCK PHOTO

Martin Elvis is a senior astrophysicist at the Center for Astrophysics | Harvard & Smithsonian in Cambridge, Massachusetts. **Charles Lawrence** is the chief scientist for astronomy and physics at NASA's Jet Propulsion Laboratory in La Cañada, California. **Sara Seager** is a physics professor at MIT in Cambridge.



Martin Elvis, Charles Lawrence, and Sara Seager

By substantially increasing the mass and volume of its reusable transportation system without raising costs, SpaceX may enable NASA to implement future missions years ahead of schedule.

From 2019 to 2021, the US astronomy community was engaged in a planning exercise for the coming decade and beyond. The result of that effort is the decadal survey *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*. Commonly known as Astro2020, it envisages an ambitious set of new "Great Observatories" as the community's top priority.[1] (Each of the authors is closely associated with one of the observatories endorsed by Astro2020.) The new Great Observatories, some of which are shown in figure 1, would collect measurements that span the electromagnetic spectrum, from far-IR to x rays, with orders-of-magnitude gain in capabilities over their renowned predecessors—the *Spitzer Space Telescope*, the *Hubble Space Telescope*, the *Compton Gamma-Ray Observatory*, and the *Chandra X-Ray Observatory*.

To keep within the NASA astrophysics budget, however, their launch dates have been pushed to the 2040s and 2050s, a forbidding timeline. A newly minted PhD today will be barely a decade from retirement by the time even the first of the observatories launches. The unwelcome implication is that there likely will be a decade-scale gap in flagship capabilities at all wavelengths in the 2030s to the detriment of science and of NASA's technological leadership.

Astro2020 took place against a rather static background of space capabilities. Yet from late 2020, SpaceX has been developing an enormous and fully reusable launch system known as Starship, which consists of the Starship upper stage and the Super Heavy booster stage. The Super Heavy hasn't flown yet, although Starship underwent dramatic progress, from early tests that resulted in multiple explosions—known tongue-in-cheek as "rapid unscheduled disassemblies"—to a successful high-altitude test flight and soft landing by mid 2021. Studies of the largest flagship missions that NASA commissioned took three years and were completed by 2019. The unfortunate timing meant that the capabilities of Starship could be only briefly considered in the Astro2020 deliberations.

Assuming it is successful, Starship will dramatically enhance our space capabilities in ways that will qualitatively alter how astrophysics missions can be built. The capabilities for planetary science missions in our solar system are discussed in the *Origins, Worlds, and Life* report, which emphasizes that Starship can accelerate the NASA planetary program.[2] This paper discusses the parallel opportunities for astrophysics.

Mass, size, and cost

Astrophysics missions to space have always been tightly constrained by the capabilities of the launchers, which have not changed substantially in two decades. The three changes that Starship would bring are a much larger mass to orbit, much wider cargo bays, and no increase in—and potentially lowering—the cost per launch.

For decades the maximum mass brought to low Earth orbit has been around 10–25 metric tons (t). The Starship Users Guide says that the spacecraft will be capable of carrying about 100 t to low Earth orbit, which is 4–10 times more than other launchers (see figure 2). Starship will be able to put 21 t into geostationary transfer orbit and about 18 t into a Sun–Earth L2 Lagrange point orbit, a favored location for many classes of astrophysics missions, including the *James Webb Space Telescope* (*JWST*). Refueling in orbit is required for NASA's lunar *Starship Human Landing System*.[3] It could transport 100 t observatories to the Moon, to the L2 orbit, or almost anywhere in the solar system.[2]

Space observatories are deployed from the cargo bay of the upper stage, known as the payload fairing, of their respective launchers. They then fly independently for their operational lives, typically years to decades. All heavy-lift vehicles launched this century have had inner fairing diameters of 4–5 m. Starship will double that diameter to 8 m and marginally increase the typical payload height, as shown in figure 3.

Most launchers cost more than $100 million to design and build. Exceptions are the Proton-M and Falcon 9 vehicles, which cost about $60 million, but the production of Proton-M vehicles ended in 2022. The goal for Starship is to be cheaper than the Falcon 9 rocket.[4] But even a launcher with zero costs would not be transformative without the large increases in payload mass and volume that Starship is designed to provide. A $60 million launch cost for NASA's Medium-Class Explorer missions, for example, is 20% of the mission's $300 million budget.

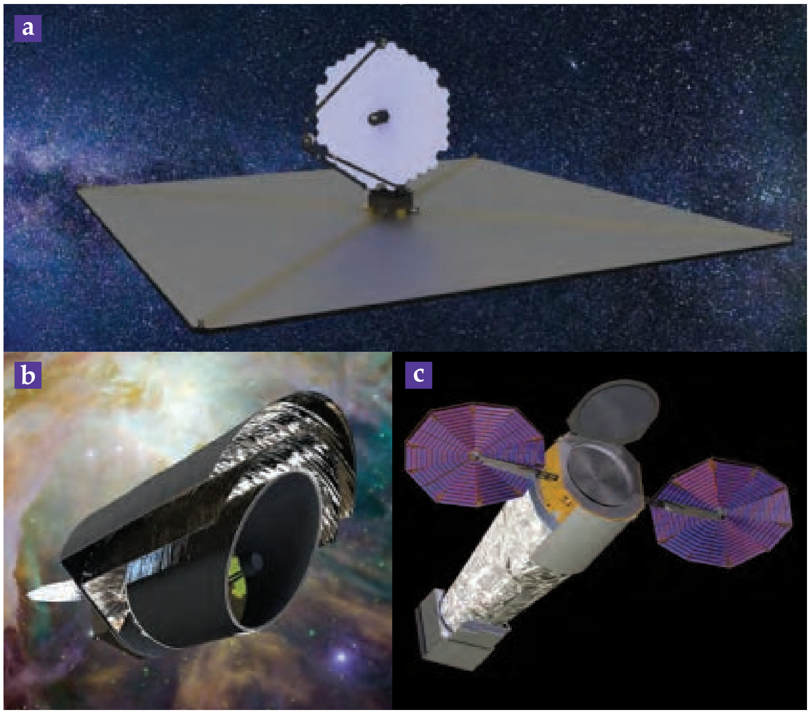


FIGURE 1. NASA is planning the development of new Great Observatories. Shown here are early design concepts: **(a)** the *LUVOIR* observatory, **(b)** *Origins Space Telescope*, **(c)** and the *Lynx X-Ray Observatory*. Whatever form the final instruments take, they will be sent to collect measurements across the electromagnetic spectrum and answer the biggest open questions in astronomy and space science. The SpaceX Starship launch vehicle could bring those observatories to space before the middle of the century. (Courtesy of NASA/GSFC.)

Revolution in mission design

Mass and volume have dominated space mission design, but Starship would reduce mass to a secondary design factor. That approach will simplify decisions and reduce the number of design cycles that must be completed before arriving at a workable solution. Spacecraft have traditionally demanded strict performance margins to save mass. But with looser mass requirements, mission designs can use simpler, heavier components and less exotic materials and incorporate more robust engineering margins. According to the *Origins, Worlds, and Life* report,

> Starships can accommodate payloads that are significantly larger and heavier than traditional NASA planetary payloads, significantly reducing the need for the costly reductions in size and mass required for traditional NASA payloads. Starships can fly multiple payloads and instruments on individual flights to reduce overall risk, and significantly more power can be available for the payload. (reference 2, page 540)

During the design phase of NASA's modest-sized *Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer* (*SPHERE*x), for example, engineers used the mass available on the SpaceX Falcon 9 launch vehicle to help solve problems and contain costs. Allen Farrington, the project manager of *SPHEREx* at NASA's Jet Propulsion Laboratory, told the three of us that "the approach that *SPHEREx* has taken from proposal through the critical design review is to convert risk to mass. A key example was the Sun–Earth shade, where we swapped out technically challenging, state-of-the-art, soft-goods-based technology for more massive but state-of-practice aluminum honeycomb panels. This resulted in a much lower risk posture and was enabled by the excess mass capabilities of our Falcon 9 launch vehicle."

The *JWST* exemplifies the difficulties caused by tight size and mass constraints. The Ariane 5 launch vehicle constrained the total payload mass to 6.2 t. The *JWST* primary mirrors, including their support structure, are ⅙ of the total mission mass. That's similar to the *Hubble* mirror but with nearly six times the area. A *Hubble*-style mirror for the *JWST* would have had a mass of almost 5 t, or ¾ of the total available payload.

The limitations of the launcher capabilities forced project scientists to develop novel, lightweight, high stiffness-to-mass technologies. Their choice of beryllium for the mirror material was driven in part by the need for high conductivity to minimize thermal gradients at the 20–55 K operating temperatures of the *JWST*.[5] The need to deploy a large, thin sunshield had other consequences, including reducing slew rates and lengthening settling times, both of which have reduced the amount of scientific work that can be done each day.[6]

Even though the *JWST* successfully deployed, vindicating the technical approach, the complexity of the design required extensive planning and testing that added to the cost and lengthened the project's schedule. With Starship's large fairing diameter and volume, the 6.5 m *JWST* primary mirror could have been made of a single component with a mass per square meter similar to *Hubble*. At 5 t, the *JWST* would still have been only 10% of the mass deliverable to the Sun–Earth L2 orbit and, therefore, not a dominant design consideration. A single mirror avoids the complexity of aligning the 18 hexagonal mirror segments. Not all such origami deployments would be avoided by using Starship; the *JWST* sunshield is still larger than Starship's proposed fairing size.

Although ambitious, reducing total mission cost by a factor of two is the crucial threshold for cost savings. The same budget can then fund twice as many missions, which would be transformative for the new Great Observatories program by potentially allowing for missions slated for the 2040s to happen in the 2030s. When a set of Great Observatories is operating contemporaneously, the pace of discoveries is accelerated because findings by one observatory often lead to new investigations by others.

Gains for all bands

Astronomy observations now are collected across more than 10 decades of frequency of the electromagnetic spectrum, from

10^8 Hz in the radio band to more than 10^{18} Hz in x rays. The ways in which missions could benefit from Starship's capabilities depend on the band. The missions proposed in Astro2020 white papers can serve as a guide. Together the missions cover virtually every band of the electromagnetic spectrum along with alternative messengers such as cosmic rays and neutrinos.

For the traditional radio band at centimeter wavelengths, the obvious next step is extending very long baseline interferometry to longer baselines than Earth's diameter. That will let researchers obtain higher angular resolution and faster and denser *uv*-plane coverage, which allows for concomitant high-dynamic-range imaging. The Russian *Spektr-R* and Japanese *Haruka* missions, with modest payloads of 1–2 t, demonstrated technical feasibility and the existence of radio sources for study.[7] Both missions produced limited results because of their moderate dish sizes of 8–10 m and their single Earth-to-space baselines.

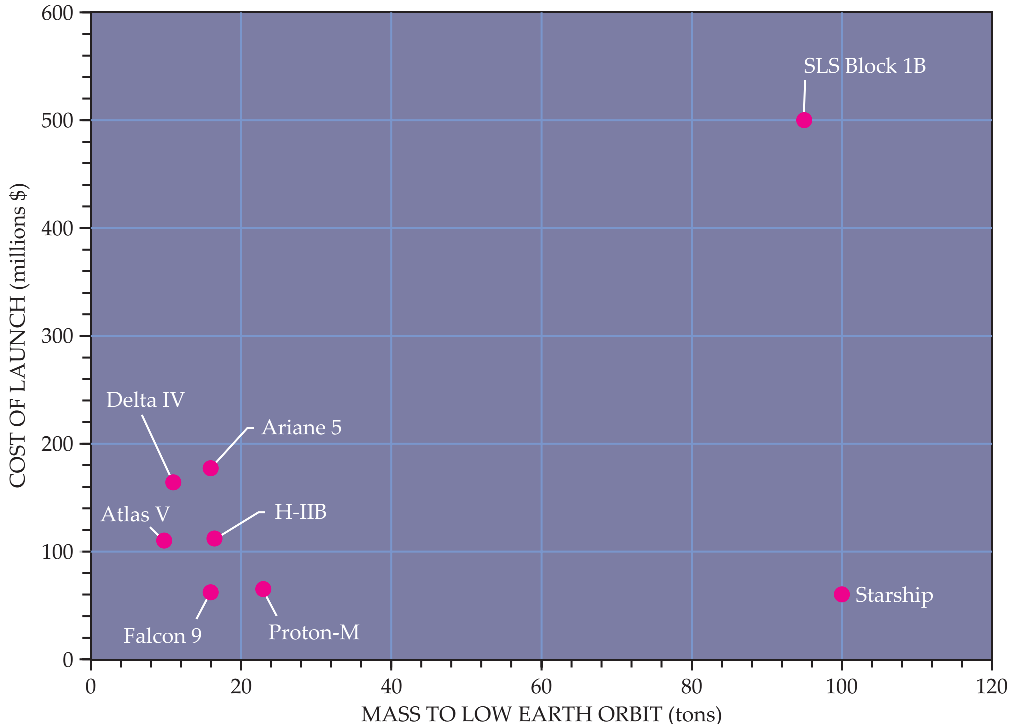


FIGURE 2. MISSIONS TO SPACE have historically been constrained by launch vehicles and the limited mass they are capable of bringing to orbit. The upcoming Starship vehicle, developed by SpaceX, could offer new opportunities by carrying more mass to low Earth orbit at a lower cost compared with the competition. (Graph by Freddie Pagani. Data are from the following: Ariane 5, Arianespace; Proton–M, International Launch Services; Atlas V and Delta IV, United Launch Alliance; H-IIB, JAXA; SLS Block 1B, NASA; and Falcon 9 and Starship, SpaceX.)

Starship could, in a single launch, deploy multiple antennae up to 30 m in diameter using a mechanism similar to the unfolding of an umbrella. The millimeter-wavelength tolerances of the antennae would allow for the detection of various celestial objects. The gain in *uv*-plane coverage from multiple antennae scales with the number of baselines $N_{BASELINES}$, which increases rapidly with the number of antennae n: $N_{BASELINES} = n(n-1)/2$. Although launch costs are unimportant with Starship, simultaneously launching an entire array could shorten project construction times and save costs.

Radio astronomy at frequencies of less than 30 MHz and wavelengths greater than 10 m could give access to the "dark ages," the time before the first stars formed, by using the cosmological signature from neutral hydrogen.[8] That approach is infeasible from Earth because of ionospheric blocking and the high human-created radio background. The lunar far side may be the only site in our solar system from which that cosmological signal is detectable because the Moon provides 90 dB suppression of Earth-based interference.[9] It is possible, though, that galactic synchrotron emission will prove to be an insurmountable source of noise. The Starship could deliver 100 t to any lunar location, so it would be able to transport a telescope, and a crew could, if necessary, reconfigure it before they deploy it.

Perhaps the most famous recent result at millimeter to submillimeter wavelengths is the 2019 image of the shadow of the supermassive black hole in the galaxy Messier 87; it came from the Event Horizon Telescope, which is based on very long baseline interferometry. The image made headlines around the world, as did the 2022 image of Sagittarius A* in the Milky Way. Theory predicts fine structure in the image, but that can't be confirmed with the 10 000 km baselines and limited frequency coverage accessible on Earth.

Longer baselines are possible from high orbits, although multiple antennae will be needed to give adequate *uv*-plane coverage. The higher angular resolution could reveal the physics at work in accelerating jets to near light speed and could increase the number of black hole shadows resolved from two to at least dozens.[10]

Really long baselines of greater than 300 000 km, made possible with the use of satellites in geostationary orbit or telescopes on the Moon, would allow a clean separation of the effects of general relativity from those of the astrophysical, even with a single baseline.[11] Optical laser-communication technology has now reached a point at which data can be transferred at a high rate from low Earth orbit, and it should be demonstrated soon on longer-baseline telescopes.

Dishes for submillimeter astronomy must be designed to tolerances of tens of microns, a constraint that makes deployed optics less appealing and increased mass more appealing. Starship could deploy a stack of several monolithic 6-m-class dishes to geostationary orbit in a single launch to help lower costs and accelerate the schedule. Another Starship could put a submillimeter telescope anywhere on the Moon.

For the far-IR region of the spectrum, the *Origins Space Telescope* is a flagship concept that has been studied by NASA. It was conceived as a 5.9-m-diameter primary mirror cooled to just 4.5 K to have low thermal background noise across the whole 25–588 micron band. *Origins* would have much more sensitivity and spectral resolution than its predecessor missions, the *Herschel Space Observatory* and *Spitzer*. Early designs for *Origins* had already planned on utilizing one of three conceptions: the larger-diameter fairings, then known as the Big Falcon rocket, from Starship; NASA's Space Launch System (SLS); or Blue Origin's planned New Glenn launcher.[12]

The *Origins* concept study did not exploit the mass capability of Starship and SLS. *Origins* has a mass of only 13 t, even if all the contingencies and reserves are included.[12] Like all far-IR observatories, *Origins* requires an orbit similar to the Sun–Earth L2 point: Such an orbit is far enough away from Earth so that its heat doesn't interfere with data collection. Even for a nonrefueled Starship, *Origins* could carry four times as much mass. That decision could lead to cost-saving opportunities from, for example, simplifying the choice of material for the primary mirror.

To support near-IR, optical, and UV astronomy, NASA conceptualized the *Habitable Exoplanet Observatory*, or *HabEx*, and *LUVOIR* missions for the Astro2020 survey.[13] Those considered projects span a wide range of possible mirror diameters, from 2.4 m to 15 m. The most demanding science goal of each concept mission is to directly image exo-Earths—the Earth-like planets in the habitable zones of the stars they orbit—and to then measure the spectrum of their atmospheres in a search for biosignatures or other signs of life. However, the stars are so much brighter than the exo-Earths orbiting them that a demanding contrast ratio of 10^{-10} is needed. To balance cost against the number of exo-Earths expected to be accessible, Astro2020 instead recommended an unnamed mission that would use a 6 m primary mirror, which is a compromise between *HabEx* and *LUVOIR*. The unnamed mission has an anticipated launch date of 2045, primarily because of its $11 billion cost estimate.

A 6 m primary mirror could be carried by Starship to orbit, and it could have *Hubble*-like mass per unit area, or even greater, without causing problems for the designers. The larger available mass that could be brought to space by Starship offers novel design possibilities for the new missions anticipated by Astro2020. Astronomers and engineers will need to explore various designs to determine whether such missions can be built at lower costs.

For x-ray observations, the *Lynx* x-ray flagship concept studied by NASA, in preparation for Astro2020, was conceived as being hundreds of times more capable of imaging and high-resolution spectroscopy than *Chandra*'s 1.2-m-diameter mirror.[14] The x-ray mirror of *Lynx* is 3 m in diameter with a subarcsecond angular resolution and features grazing-incidence optics, which reflect x rays at shallow angles. The mirror assembly constitutes 25% of *Lynx*'s mass, taking up 2 t of the envisaged 7.7 t total. To keep the mirror mass from exceeding the pre-Starship available payload, the *Lynx* scientists opted for thin, 0.5 mm mirror segments for its grazing-incidence optics. Starship would allow an x-ray mirror made of thicker, 2 mm segments. Since stiffness varies with thickness cubed, the mirror segments would be 60 times as stiff.

Attaining the subarcsecond image quality at the heart of the *Lynx* mission would then be much more easily accomplished. A simpler fixturing and alignment system that could be more rapidly assembled would likely lead to cost savings. As on other missions, the freeing of the mass constraint may lead to a lower-cost payload and spacecraft. The resulting 8 t mirror assembly is readily accommodated by Starship.

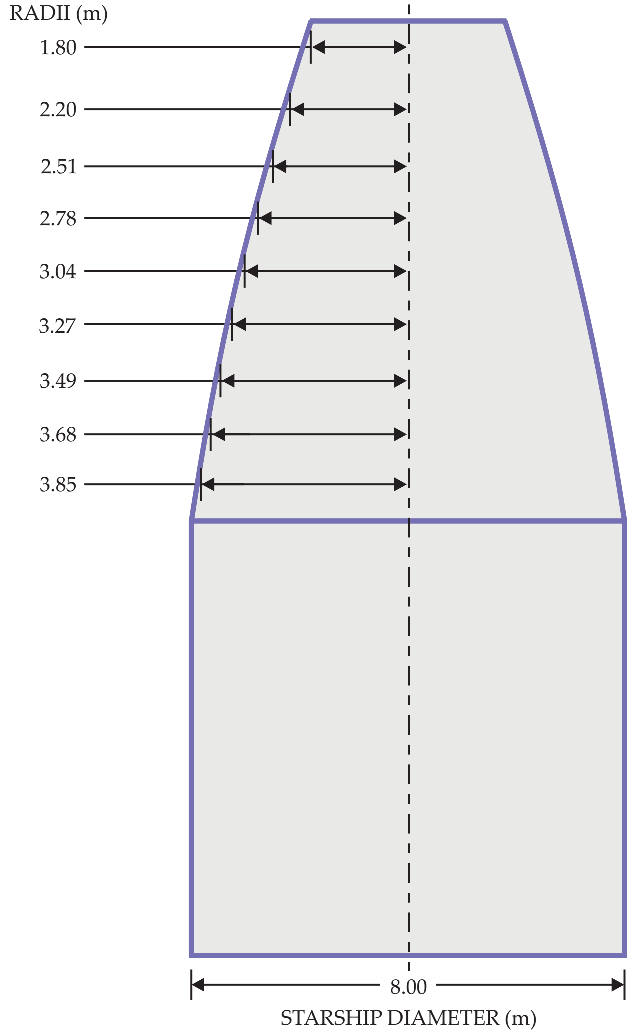


FIGURE 3. SPACEX'S STARSHIP would redefine launcher capabilities by offering twice-as-wide payload fairings, which would allow for more massive and less complex instrumentation to be brought to space. Starship's launcher length of 17.24 m is a bit more accommodating than the industry's more typical launcher lengths of 15–16 m. (Courtesy of SpaceX.)

Outside of the new Great Observatories, there is a host of novel ideas for more modest-scale instruments. The *Probe of Extreme Multi-Messenger Astrophysics* (*POEMMA*) mission,[15] for example, would use a pair of 4 m Schmidt telescopes pointing down from orbit to image extensive air showers (EASs) in stereo using fluorescence and Cherenkov flashes. EASs are created by both ultra-high-energy cosmic rays and neutrinos in Earth's atmosphere. *POEMMA* images such cosmic rays when it is nadir-pointing and neutrinos as upward-moving EASs when it is limb-pointing. The large atmospheric volume probed from orbit gives *POEMMA* a 10- to 100-fold performance gain over state-of-the-art telescopes.

The low cost of Starship would allow for the launch of two *POEMMA* telescopes separately, and the launch vehicle's large volume would remove the need for deployment mechanisms. An improvement of up to a factor of three in the collecting area could be gained by using the wide Starship fairing to launch larger, 6–7 m telescopes, although the cost of manufacturing the necessary 6 m Schmidt corrector lenses may preclude that option.

POEMMA is just one of the many probe-class mission concepts submitted in white papers for Astro2020. The $1.5 billion cost for probes estimated by Astro2020 means that only one per decade is affordable for the NASA astrophysics budget. Starship may enable cheaper probes so that more, and more unconventional ones such as *POEMMA*, can be developed.

Cheaper, faster, but beware of better

The space-science community can accelerate the Astro2020 program by taking advantage of Starship's potential for cost savings, but that approach will require discipline from all involved. "Faster, better, cheaper" was the mantra of Daniel Goldin, the NASA administrator from 1992 to 2001, and it led to, at best, mixed results.[16]

Starship seems poised to provide faster and cheaper launch vehicles. The teams proposing missions will always want to put all the available mass budget, however large, into bigger mirrors and more instruments. That line of thinking leads, in many cases, to large and complex designs that will follow the expensive scaling of cost with mass that the astronomy community is used to. Pushing for "better" could jeopardize the faster and cheaper goals, so the community will need to develop best practices to restrain scientists' appetite.

Space agencies will need to monitor for and avoid mission creep, but doing so will not be easy. Industry and agency models that predict mission cost often scale cost with mass. Starship could usher in a new paradigm in which increased mass would decrease cost. But that won't be an easy exercise. Because there is no track record showing whether an approach that uses mass and volume to cut costs will be successful, that approach translates initially into higher risk.

Starship caveats

Starship may not reach expectations. It may operate, but at a much higher cost and at a reduced mass capability, or on-orbit refueling may not be achieved. The Starship launch costs given by SpaceX are presumably estimates of the cost to SpaceX, not the price to a customer, which will be more expensive. Perhaps most importantly, realizing dramatically lower launch costs depends on rapid and frequent reuse of each Starship, but a market for sufficient launches may not be forthcoming. The anticipated savings promised by Starship also may prove illusory after careful inspection.

Similar risks, however, apply to almost any new technological development. They are thus insufficient reasons to not consider what might have the biggest effect on astrophysics if the Starship technology is a success.

The NASA-developed SLS has comparable capabilities to Starship in terms of mass to low Earth orbit and payload volume. As such, it provides some backup for Starship. The $800 million to $2.7 billion cost estimate of an SLS launch, however, would be a major factor in any mission of even a $5 billion Great Observatory.[17] Launch costs of that magnitude may put such an astrophysics mission out of contention, unless politically mandated. The SLS is fully expendable, so the rate of production of more launchers is a critical consideration. The production rate for Boeing, the lead contractor for the SLS, is limited to at most two SLS launchers per year.[18] NASA's Artemis human spaceflight program is expected to take most of the SLS launch slots over the next several years.[17] Could three launches over the next decade or so be available for the new Great Observatories?

Even if Starship works as advertised, extra mass is not without disadvantages. More mass increases the moment of inertia of the spacecraft and so requires more massive reaction wheels to point to a target. In addition, station keeping in the popular Sun–Earth L2 halo orbits either will require proportionately more propellant or will limit mission lifetimes because of the extra mass.

Starship will likely be proven or not within the next five years. That gives NASA time to prepare for a new era of launch capability by the Astro2020 midterm review. A series of coordinated studies over the next few years to investigate in detail how Starship might accomplish, accelerate, and expand the Astro2020 program would prepare NASA's astrophysics program to act if Starship succeeds. But even if Starship fails, the effort that is lost by planning for its success is small when compared with the potential gains to astronomy.

The authors thank Lee Armus, Jack Burns, Allen Farrington, Tom Megeath, Joe Silk, and Alexey Vikhlinin for valuable conversations. The cost information contained in this article is of a budgetary and planning nature and is intended for informational purposes only. It does not constitute a commitment on the part of the Jet Propulsion Laboratory and Caltech.

REFERENCES

1. National Academies of Sciences, Engineering, and Medicine, *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*, National Academies Press (2021).
2. National Academies of Sciences, Engineering, and Medicine, *Origins, Worlds, and Life: A Decadal Strategy for Planetary Science and Astrobiology 2023–2032*, National Academies Press (2022).
3. T. Burghardt, "After NASA taps SpaceX's Starship for first Artemis landings, agency looks to on-ramp future vehicles," NASASpaceflight.com, 20 April 2021.
4. A. Mann, "SpaceX now dominates rocket flight, bringing big benefits—and risks—to NASA," *Science*, 20 May 2020.
5. H. Warren, "The mirror of the *James Webb Space Telescope*: Looking into the past," NASASpaceflight.com, 13 November 2021.
6. K. Q. Ha, M. D. Femiano, G. E. Mosier, in *Proceedings of SPIE, Volume 5528: Space Systems Engineering and Optical Alignment Mechanisms*, L. D. Peterson, R. C. Guyer, eds., SPIE (2004), doi:10.1117/12.558456.
7. Y. Y. Kovalev et al., in *XXXIth URSI General Assembly and Scientific Symposium*, IEEE (2014), doi:10.1109/URSIGASS.2014.6929994; H. Hirabayashi et al., *AIP Conf. Proc.*, **599**, 646 (2001).
8. J. D. Bowman et al., *Nature* **555**, 67 (2018); J. Silk, *Found. Phys*. **48**, 1305 (2018).
9. J. O. Burns et al., *Planet. Sci. J.* **2**, 44 (2021).
10. D. W. Pesce et al., *Astrophys. J.* **923**, 260 (2021).
11. M. D. Johnson et al., *Sci. Adv.* **6**, eaaz1310 (2020).
12. M. Meixner et al., https://arxiv.org/abs/1912.06213.
13. S. Gaudi et al., *Bull. Am. Astron. Soc.* **51**(7) (2019), paper no. 89; *LUVOIR* team, https://arxiv.org/abs/1912.06219.
14. A. Vikhlinin, *Bull. Am. Astron. Soc.* **51**(7) (2019), paper no. 30.
15. A. V. Olinto et al., *Bull. Am. Astron. Soc.* **51**(7) (2019), paper no. 99.
16. H. E. McCurdy, *Faster, Better, Cheaper: Low-Cost Innovation in the U.S. Space Program*, Johns Hopkins U. Press (2001).
17. E. Berger, "Finally, we know production costs for SLS and Orion, and they're wild," *Ars Technica*, 1 March 2022.
18. J. Foust, "NASA and Boeing look ahead to long-term SLS production," *SpaceNews*, 11 December 2019. PT

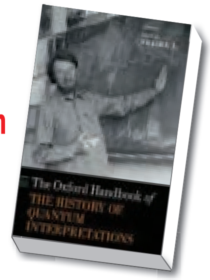
PAUL EHRENFEST JR/AIP ESVA/WEISSKOPF COLLECTION

WERNER HEISENBERG (left) and Niels Bohr (right) in heated conversation during a meal at a 1936 conference at the Institute for Theoretical Physics (now the Niels Bohr Institute) at the University of Copenhagen.

Questioning quantum orthodoxy

The Oxford Handbook of the History of Quantum Interpretations

Olival Freire Jr, ed.
Oxford U. Press, 2022. $190.00



One might have hoped that the 2022 Nobel Prize in Physics, which was awarded to Alain Aspect, John Clauser, and Anton Zeilinger for enormous contributions to the foundations of quantum mechanics, would finally put to rest the long-standing controversies over that theory's interpretations. If anything, though, the prizewinning work only emphasized the aggravating features that have driven those debates over the last century.

The editorial team behind the new *Oxford Handbook of the History of Quantum Interpretations* justifies the need for its monumental volume by foregrounding that paradox. As noted in the introduction, quantum mechanics is a fantastically successful theory that remains radically ambiguous in its foundations and meaning. It demands interpretation but remains strangely recalcitrant. There are many competing interpretations, all of which are unsatisfactory in some way, and it seems there will be no resolution anytime soon.

The editors, who are led by the historian of science Olival Freire, argue that a handbook is the best way to tackle such a controversial subject because any single-authored volume would inevitably be biased. Instead we have a wonderfully diverse group of authors, disciplinary perspectives, writing styles, and levels of technicality. The downside of the format is its sheer scale: 51 full-length articles plus an introduction, which span over 1200 pages. The diversity is certainly an asset but can sometimes be confusing. For example, there are easily a dozen different presentations of Niels Bohr's philosophy of physics, which are all intellectually stimulating but perhaps overwhelming for someone new to the material.

But anyone with a grounding in quantum physics or its history and philosophy will find the handbook an extremely valuable volume. Although it can function as an encyclopedia (What are the differences between de Broglie–Bohm theory and modal interpretations, again?), it strikes me more as something of a quantum superposition of a few books. It could be a history of quantum physics. But depending on how you read it, it could also be a book on the historiography of quantum physics, or one on the measurement problem, or one on the sociocultural influences on modern physics, or one on what it means to interpret a theory at all.

The connections between the articles make for fascinating reading. One can learn about the famed debates over quantum theory that occurred at the early Solvay Conferences on Physics through José Perillán's chapter on rhetorical strategies used to support various interpretations, or through Richard Staley's look at how concepts formed. Or one could examine the role of experiments in quantum interpretations through Climério Paulo da Silva Neto's chapter on tabletop instruments, or through David Kaiser's contribution on probing cosmological entanglements. I imagine the best audience for this book is someone like myself, who is familiar with one aspect of quantum scholarship and wants to get up to speed on another. Anyone thinking of starting a research project touching on quantum interpretations should start here.

One of the many interesting themes that emerge across the volume is how it prompts the reader not only to question the orthodox understanding of quantum mechanics—namely, the Copenhagen interpretation—but to question the very idea that there ever was such a quantum orthodoxy. In other words, the assumption that there was or is a default approach to quantum mechanics comes under a great deal of scrutiny.

Don Howard's chapter leads the charge in dismantling the Copenhagen myth, but many of the other authors also contribute to complicating its traditional position in the historiography. They show, for example, that the interpretation was not stable over time (it differed before and after World War II), over place (opinions on what the interpretation was diverged in Denmark and Germany), or even within the Copenhagen social circle itself (Bohr and Werner Heisenberg never really agreed on anything).

That creates a sense of a genuinely open field: All the modern interpretations are on something of an equal footing. There is not really a reigning champion that needs to be dethroned. Instead, the approaches can demonstrate their strengths and weaknesses on their own terms. Bohr's ideas are not necessarily the point of departure.

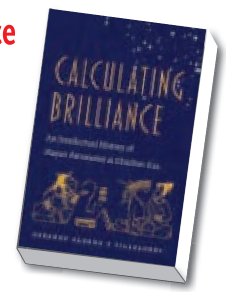
That said, having so many plausible contenders in one place, each presented so forcefully, can be somewhat dizzying. But it surely reinforces the initial premise of the volume: that the interpretation of quantum mechanics is a live field, one with rich opportunities for innovation and creative thinking. It also emphasizes the strange nature of quantum theory, which is remarkably productive despite seeming unstable and fractured. Its mysteries are clearly not going to be resolved in the near future, which makes this handbook a launching point for future work rather than a summary of past efforts.

Matthew Stanley
New York University
New York City

The famous step pyramid at the center of Chich'en Itza (background) was dedicated to the feathered serpent deity (foreground).

Understanding the Mayans on their own terms

Calculating Brilliance
An Intellectual History of Mayan Astronomy at Chich'en Itza
Gerardo Aldana y Villalobos
U. Arizona Press, 2022. $75.00



The works of Mayan astronomers have long been an object of popular fascination. Their impressive feats include a base-20 system for calculating large numbers, which included the number zero and negative quantities; a 365-day calendar system based on an accurate measurement of the solar year that the Mayans diligently corrected on a 52-year cycle; and another 260-day calendar created for ritual and divination purposes. They used the two calendars to create their famous long count, a system that numbered every day from some remote, mythical origin date. All those achievements were made before 900 CE, when Mayan civilization in major towns collapsed.

As Gerardo Aldana y Villalobos, a historian of science, details in his new book, *Calculating Brilliance: An Intellectual History of Mayan Astronomy at Chich'en Itza*, Western scholars long focused their attention on those intricate Mayan calendar systems because they aimed to determine how the Mayan system corresponded with the Western calendar. That type of Eurocentrism in Mayan studies, he argues, has led to a flawed understanding of Mayan astronomy that removes it from its cultural context.

In *Calculating Brilliance*, Aldana aims to rectify that situation by outlining the history of our understanding of Mayan astronomy and presenting a daring conclusion of his own: that a female astronomer named K'uk'ul Ek' Tuyilaj who worked in the city of Chich'en Itza likely obtained data relating to Venus's orbital trajectory.

As Aldana details, the looting and dispersion of Mayan documents following the Spanish conquest of Mesoamerica delayed the systematic study of Mayan science and society for centuries. It was only at the end of the 19th century that the German librarian Ernst Förstemann made a foundational contribution to the understanding of Mayan culture by deciphering part of the Dresden Codex, possibly the most significant collection of Mayan hieroglyphic texts. Förstemann was able to unravel the Mayan number system and tried to decode a portion of the codex that contained what appeared to be records of appearances of Venus—which is now known as the Venus table.

The 20th century saw many breakthroughs in the study of Mayan astronomy, including a better appreciation of the relationship between Mayan architecture, inscriptions, and astronomical observations. During that time, however, few scholars attempted to decipher Mayan hieroglyphic writings and inscriptions, which prevented them from analyzing astronomy in the context of broader Mayan society. It was only in the second half of the 20th century that the Russian American archaeologist Tatiana Proskouriakoff decoded the hieroglyphs and revealed what the inscriptions on Mayan monuments tell us about historical events.

That was when Aldana, who teaches at the University of California, Santa Barbara, entered the field. He began working on Mayan astronomy during his doctoral studies at Harvard University under Owen Gingerich. At that time, he also began engaging with historical and archaeological work on Mayan history and archaeology. In *Calculating Brilliance*, Aldana connects the Dresden Codex's Venus table—which is now more fully understood because of Proskouriakoff's work deciphering the hieroglyphs—with the inscriptions on Chich'en Itza's buildings, where a mural prominently depicts K'uk'ul Ek' Tuyilaj.

Following the work of Anthony Aveni, who in the 1980s suggested that the observations for the Dresden Codex's Venus table were made in a structure at Chich'en Itza called the Caracol, Aldana argues that K'uk'ul Ek' Tuyilaj likely made corrections to the table to account for data she and other astronomers had collected on the planet's orbital trajectory. Those corrections were intended to cope with the small difference between Venus's 584-day orbit and what we call the synodic period of Venus. That allowed her to predict how morning appearances of Venus would coincide with certain dates from the 260-day divinatory calendar.

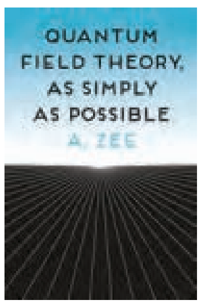
But unlike prior scholars, Aldana isn't interested in assessing the accuracy of Mayan astronomy according to contemporary Western standards. Using historical and anthropological records, he aims to illuminate the role astronomy played in Mayan society. As he emphasizes, K'uk'ul Ek' Tuyilaj's discoveries not only helped build calendars but were integral to ceremonial and ritual events. Aldana also notes that the Mayans mastered the Venus table around the same time that the feathered serpent, a god worshipped by other Mesoamerican societies, became a popular deity for Mayans to revere. That may have led the Mayans to identify the serpent with Venus.

Unfortunately, the print version of the book suffers from poor-quality images, which look much better in the digital version. Nevertheless, *Calculating Brilliance* is a tribute to humanity's cultural diversity. The book marks a milestone in our understanding of Mayan astronomy and culture by illustrating how they were inherently intertwined.

Olival Freire Jr
Federal University of Bahia
Salvador, Brazil

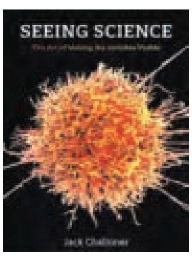
NEW BOOKS & MEDIA

Quantum Field Theory, as Simply as Possible



A. Zee
Princeton U. Press, 2023. $39.95

Countless books aim to explain quantum mechanics to a lay audience, but few authors have attempted to present a popular introduction to quantum field theory, the mathematical framework linking quantum mechanics with special relativity. A. Zee, a well-known author of several textbooks and popular-science works, attempts to fill that gap in this new book. But as he freely admits, it isn't for the faint of heart: Be prepared to see a fair share of summation symbols, path integrals, and Greek letters. If readers are willing to put in a bit of work, they will be rewarded. Zee's witty, insightful writing and engaging historical anecdotes make the book a pleasure to read. —RD



Seeing Science

The Art of Making the Invisible Visible

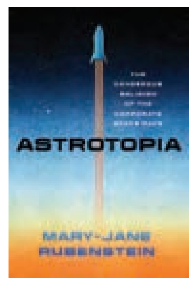
Jack Challoner
MIT Press, 2022. $34.95 (paper)

In this new coffee-table book, the science writer Jack Challoner has collected and annotated more than 200 images—historical and contemporary—of phenomena like neurons, animal locomotion, and the Sun's photosphere. But *Seeing Science* isn't just a collection of beautiful pictures. Challoner also aims to answer a question that many laypeople often wonder about: How do scientists produce images of objects or phenomena they can't see with the naked eye? To do so, he intersperses the text with explanations of how such technologies as scanning electron microscopes and space telescopes have enabled researchers to image seemingly unseeable things. The result is a compelling collection. —RD

Astrotopia

The Dangerous Religion of the Corporate Space Race

Mary-Jane Rubenstein
U. Chicago Press, 2022. $24.00



In *Astrotopia*, the philosopher Mary-Jane Rubenstein argues that the 21st-century private space race being carried out by Elon Musk, Jeff Bezos, Richard Branson, and others has become a "mythological project" analogous to the type of "imperial Christianity" that was used by Europeans to colonize more than half the planet. Discussing the era of private competition in space, known as NewSpace, she argues that we need to act now to prevent it from being rapaciously exploited by capitalists. Yet perhaps the most provocative portion of the book looks toward the past: Rubenstein convincingly demonstrates that NASA and US politicians used Christian imperialist language to justify the Apollo missions. In other words, it's no coincidence that the *Apollo 8* crew read from the book of Genesis while orbiting the Moon on Christmas Eve 1968. —RD PT

NEW PRODUCTS

Focus on test, measurement, quantum metrology, and analytical equipment

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

Software for quantum effects control

Quantum Machines has enhanced the capabilities of its Quantum Orchestration Platform (QOP), which integrates parametric pulse programming, real-time classical processing, flow control, and ultrafast analog feedback. New features in release 2.0 improve fidelity, simplify setups, and provide greater flexibility to run advanced quantum experiments. The intuitive, open-source QUA programming language lets users run advanced quantum algorithms out of the box. QOP V2.0 offers strict timing for sequences with zero gaps, a crosstalk matrix, improved digital filters, advanced chirping, an input stream for fast data transfer during run time, high-resolution time tagging, and shareable ports. ***Quantum Machines***, *Yigal Alon St 126, Tel Aviv-Yafo, Israel, www.quantum-machines.co*

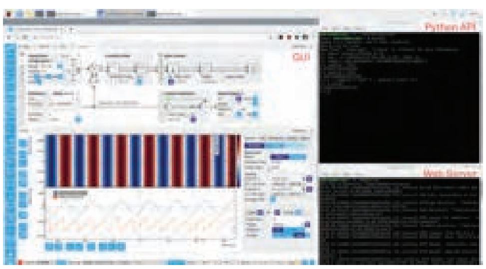
Digital pressure gauge

Omega's latest series of digital pressure gauges, DPG509, offers full-scale accuracy as high as 0.08%; 0.25% accuracy is also available. The DPG509 technology builds upon a diffused silicon piezoresistive sensor to create an efficient digital platform that maintains high accuracy, improves temperature stability, and enables compensation in both the positive and the negative pressure direction. With the data-logging option, users can log data as fast as four readings/s. Data analysis is straightforward with an included, easily accessible SD card. All options have a thermoplastic housing with an ingress-protection rating of IP65 for durability in harsh environments. The highly configurable DPG509 series can be used in a wide range of applications, including testing, hydraulics and pneumatics, pumps and compressors, fluid management, level indication, filter monitoring, and automation. ***Omega Engineering Inc***, *800 Connecticut Ave, Ste 5N01, Norwalk, CT 06854, www.omega.com*

Fast oscilloscope

Rohde & Schwarz has added a new series, the R&S MXO 4, to its oscilloscope portfolio. Because of an application-specific integrated circuit with a data processing rate of 200 GB/s, the R&S MXO 4 series can achieve the world's fastest real-time update rate of more than 4.5 million acquisitions per second, according to the company. A 12-bit analog-to-digital converter operates across all the instrument sample rates with an 18-bit vertical-resolution architecture. The R&S MXO 4 series also features a standard acquisition memory of 400 megapoints on all four channels and the lowest noise and largest offset range in its class: ±5 V with a scaling of 500 μV/div. Once only available in more costly higher-performance oscilloscopes, the digital trigger comes standard in the new series. The trigger sensitivity of 1/10 000 division enables users to isolate difficult-to-find small physical layer anomalies in the presence of large signals. The oscilloscopes in the R&S MXO 4 series are the first to perform 45 000 FFTs/s. ***Rohde & Schwarz GmbH & Co KG***, *Mühldorfstraße 15, 81671 Munich, Germany, www.rohde-schwarz.com*

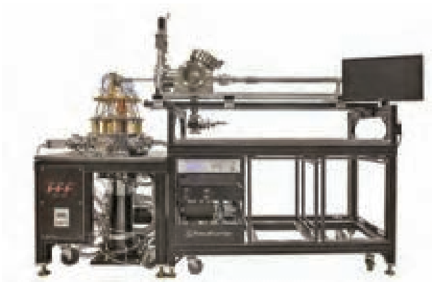
Software for scalable quantum computing



Zurich Instruments designed its LabOne Q software for experiments on advanced control electronics. The Python-based interface provides intuitive tools for specifying multichannel signals with precise timing control. LabOne Q users can program large setups in Python as a single machine and transform complex quantum circuits into simple code, which can easily be reused. Using sections, users can group pulse patterns across signal lines. To express all levels of timing consistently, the sections support looping, sweeping, and averaging structures. Although users have complete low-level access to the instruments, the software streamlines and automates time-consuming tasks such as generating and uploading waveforms, synchronizing pulses between multiple instruments, and optimizing instrument settings, which can help maximize computational duty cycles. Upgrading or scaling up a laboratory setup requires only a drop-in replacement in the code. ***Zurich Instruments AG***, *Technoparkstrasse 1, 8005 Zürich, Switzerland, www.zhinst.com*

Arbitrary waveform generator

According to Keysight, its M8199B is the first arbitrary waveform generator (AWG) with a sampling rate up to 256 GS/s that offers analog bandwidth exceeding 80 GHz and up to eight synchronized channels operating simultaneously. It enables data transmission greater than 400 GB/s per lane in intensity-modulation and direct-detect formats and greater than 1.6 TB/s per carrier in coherent optical communications. The M8199B AWG is suitable for high-speed research experiments using multilevel pulse- and quadrature-amplitude and other proprietary modulation formats at symbol rates approaching 200 GBaud. Users conducting physics, chemistry, and general-purpose electronics research can generate any arbitrary waveform that can be mathematically described, including the ultrashort pulses, wideband RF pulses, and chirps needed for applications such as chemical reactions, elementary-particle excitation, and quantum effects. ***Keysight Technologies Inc***, *1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*



Quantum device testing

FormFactor has unveiled its IQ2000 die probing system that can cool devices from room temperature to a base temperature of 4 K in less than 1 h and simultaneously enable probe testing. It can expedite testing for quantum control chips, cryo-CMOS, photonic devices, niobium-based circuits, materials development, and other applications that require ultralow-temperature environments. The IQ2000 load-lock chamber and cryogenic probe head provide tool-free device exchange without opening or warming up the ultracold test environment. The high-density electrical probe interface enables high-bandwidth, parallel device contact. The IQ2000 offers options for base temperatures of less than 2 K or 4 K. It features up to 128 low-frequency signals and up to 28 high-bandwidth (greater than 12.5 GHz) signals for parallel device testing with FormFactor's probe-card technologies. The low-vibration system provides for stable device contact and low-noise measurements. ***FormFactor Inc***, *7005 Southfront Rd, Livermore, CA 94551, www.formfactor.com*

Multichannel coherent microwave RF source

Rigol now offers an RF microwave generator, the DSG5000 series, for complex multichannel and system-level applications. The new signal generators are available with 2, 4, 6, or 8 independent RF channels. Designed for high-frequency signals in applications such as quantum research and radar, the DSG5000 series features high long-term phase stability between channels of ±1°. At 10 GHz with a temperature variation of less than 1 °C, the phase stability can be measured at less than 1° of phase deviation over 72 h. According to the company, that stability yields the highest accuracy in long-term verification applications compared with competitive solutions. The RF source can generate up to 20 GHz signals with amplitude-, frequency-, phase-, and pulse-modulation capabilities. It provides high signal purity with low phase noise measuring less than −33 dBc/Hz at 1 GHz, an offset of 10 kHz, a fast switching speed of 3 ms, and a typical amplitude accuracy of ± 1.1 dB. ***Rigol Technologies Inc***, *10220 SW Nimbus Ave, Ste K-7, Portland, OR 97223, www.rigolna.com*

High-resolution SWIR cameras for thermography

InfraTec developed its ImageIR 8100 and ImageIR 9100 short-wave IR (SWIR) cameras to offer high measurement accuracy and very good short- and long-term stability in thermographic measurement tasks involving very high temperatures and challenging materials. For example, the cameras allow emissivity-optimized measurements of temperatures on metallic surfaces in applications between 300 °C and 1700 °C. They are also suitable for laser applications, additive manufacturing, and welding processes and press hardening in the metal industry. The ImageIR 8100 and ImageIR 9100 cameras are radiometrically calibrated in the VGA and SXGA image formats with 640 × 512 and 1280 × 1024 IR pixels, respectively. Both have a 5 μm pixel pitch, which allows for a comparatively affordable, compact optical design with high imaging quality. Brilliant thermographic images with high geometric and thermal resolution can be achieved in both formats. ***InfraTec Infrared LLC***, *5048 Tennyson Pkwy, Ste 250, Plano, TX 75024, www.infratec.eu* PT

OBITUARIES

Kurt Gottfried

Kurt Gottfried, a professor of physics at Cornell University who was also known for his work to promote nuclear arms control, secure the human rights of scientists, and prevent political interference in science, died on 25 August 2022 in Ithaca, New York.

Born on 17 May 1929 in Vienna, Kurt fled with his family to Belgium in 1938 after *Kristallnacht*, when the Nazis carried out raids against the Jewish population. The Gottfried family ultimately emigrated to Montreal in 1939.

After receiving his bachelor's and master's degrees in engineering physics at McGill University in 1951 and 1953, respectively, Kurt earned a PhD in theoretical physics at MIT in 1955. His thesis, "Investigations based on the Bohr–Mottelson nuclear model," was done with Victor Weisskopf as his adviser. Kurt spent most of the next nine years at Harvard University, as a junior fellow and then an assistant professor, with a short stint at the Niels Bohr Institute, before joining the Cornell physics department in 1964. He became a professor emeritus in 1998.

Kurt is known for his early work in 1963–64 with John David Jackson in analyzing long-lived unstable particles from high-energy hadronic collisions in order to understand the strong interaction. He published his widely used textbook, *Quantum Mechanics: Fundamentals*, in 1966. He published a second, expanded edition, coauthored with one of us (Yan), in 2003.

Kurt also made significant contributions to the understanding of quarks. In 1967, as a test of the quark model, he developed the Gottfried sum rule, which quantifies the difference in the electron scattering probability off different nucleons according to their quark composition.

In 1975 Kurt and Cornell colleagues (including Yan) explained the recently discovered J/ψ and ψ' particles as bound states of the predicted charm quark and antiquark pair (charmonium) and calculated the energy spectrum of those states. Experimental confirmation of their predictions was important in demonstrating the existence of quarks.

Kurt spent 1968–69 at MIT. Student activism was high, especially around the involvement of academic science departments in projects related to the Vietnam War and new weapons systems. Kurt drafted a statement urging scientists and engineers to examine society's use of science and technology and calling for a day of education on 4 March 1969, during which university classes would devote their time to discussing those issues. Several dozen colleges across the country took part.

With Herman Feshbach and Francis Low, Kurt led a newly formed faculty group—the Union of Concerned Scientists (UCS)—which released the statement, signed by a list of prominent scientists, and organized two days of speakers at MIT. In the 1970s Kurt and Henry Kendall turned the UCS into an independent nonprofit organization (today with more than 200 staff), where scientists and other experts could apply their knowledge to important public policy issues. Kurt served on the UCS board until his death, and he was its chair from 1999 to 2009.

During the 1980s and 1990s, Kurt spoke and wrote on a range of security issues, contributing to the *New York Times*, the *Washington Post*, *Scientific American*, and other outlets. In 1982–83 he, Richard Garwin, and Leonard Meeker developed a draft treaty banning anti-satellite and space-based weapons, which they presented to Congress. Following President Ronald Reagan's 1983 announcement of the Star Wars program, Kurt worked closely with prominent scientists, including Kendall, Garwin, and Hans Bethe, and with UCS staff to critique the missile defense program on technical and arms-control grounds. He coauthored the influential book *The Fallacy of Star Wars* (1984). With two of us (Gronlund and Wright), Garwin, and others, he wrote *Countermeasures: A Technical Evaluation of the Operational Effectiveness of the Planned US National Missile Defense System* (2000).

Expanding his focus to include nuclear weapons command and control, Kurt coedited *Crisis Stability and Nuclear War* (1988) with Bruce Blair. He also worked with Paul Bracken on post–Cold War security relations, coediting *Reforging European Security: From Confrontation to Cooperation* (1990).

During the 1980s Kurt traveled to the Soviet Union to meet with refuseniks and urged other scientists to actively support the dissidents. He was instrumental in securing the release of the prominent Soviet physicist Yuri Orlov in 1986 and arranged for him to join the Cornell physics faculty. Kurt also helped obtain the release in 1978 of the Argentine physicist Elena Sevilla, imprisoned for her husband's political activities, and was able to get her to Cornell to finish her doctorate degree.

Kurt Gottfried

In 2004, in response to the manipulation of the scientific advisory process by the George W. Bush administration, Kurt worked with the UCS to draft the statement "Restoring Scientific Integrity in Policy Making" and recruit 62 preeminent scientists to sign. Ultimately endorsed by 12 000 scientists, it drew public attention to the issue and led to strengthened scientific-integrity guidelines and policies throughout the federal government.

For his lifetime of work, the American Association for the Advancement of Science awarded Kurt its 2016 Scientific Freedom and Responsibility Award.

Kurt was humble, was respectful of others, and had a droll sense of humor. His dedication to making positive change in the world was an inspiration to many scientists. He thought deeply about issues, often working in partnership with his wife, Sorel Gottfried. UCS board members aptly called him the organization's "moral compass."

Lisbeth Gronlund
David Wright
Massachusetts Institute of Technology
Cambridge
Tung-Mow Yan
Cornell University
Ithaca, New York

Alvin William Trivelpiece

An outstanding leader, physicist, and administrator, Alvin William Trivelpiece died on 7 August 2022 in Rancho Santa Margarita, California.

Born on 15 March 1931 in Stockton, California, Al put himself through California Polytechnic State College and graduated in 1953 with a degree in electrical engineering. He earned a PhD in electrical engineering in 1958 from Caltech under the guidance of Roy Gould. His dissertation was titled "Space charge waves in cylindrical plasma columns."

After spending a year in the Netherlands as a Fulbright scholar, Al became a professor of electrical engineering at the University of California, Berkeley. In 1966 he became a professor of physics at the University of Maryland, College Park, where he supervised 19 doctoral candidates and published more than 40 papers on plasma physics. Al also wrote, with Nicholas Krall, *Principles of Plasma Physics*, which remains a classic textbook 50 years after publication.

Taking a leave of absence from the University of Maryland in 1973, Al served for two years with the Atomic Energy Commission as an assistant director for research in the division of controlled thermonuclear research. Among his achievements was the establishment of the national Controlled Thermonuclear Research Computer Center (CTRCC), which provided fusion researchers across the country with remote access to leading-edge computing resources. The idea was met with considerable skepticism, but it became a model used by other science agencies.

In 1976 Al became vice president for engineering and research at Maxwell Laboratories, and in 1978 he was appointed corporate vice president of Science Applications Inc. Three years later Al was nominated by President Ronald Reagan and confirmed by the Senate as director of the Department of Energy's Office of Energy Research (now the Office of Science), where he had responsibility for 5 research programs and 10 of DOE's laboratories. He also served as the energy secretary's science adviser.

As director, Al developed a plan for construction of five major scientific research facilities: the Advanced Light Source, the Advanced Photon Source, the Relativistic Heavy Ion Collider, the Continuous Electron Beam Accelerator Facility, and the Spallation Neutron Source (SNS). Those facilities continue to sustain US leadership in many fields of science today. Al's planning was informed by federal advisory committees that he established and the National Academy of Sciences, an approach that continues to serve DOE's Office of Science well.

Al and Charles DeLisi in 1986 initiated the Human Genome Project, which has had an enormous impact on the life sciences and medical research. Al was masterful in obtaining support for the project from DOE management, the Office of Management and Budget, and the US Congress. The National Institutes of Health joined the project two years later.

Al expanded the scope of the CTRCC to form the National Energy Research Scientific Computing Center, which continues to serve the research programs in the Office of Science as well as other DOE programs.

In 1987 Al became the executive officer of the American Association for the Advancement of Science and publisher of *Science*. Then in January 1989, he became director of Oak Ridge National Laboratory, where he was responsible for several initiatives. He created the Office of Laboratory Computing to enhance Oak Ridge's ability to meet complex scientific challenges. Al led the development of a successful proposal for the Center for Computational Sciences, which brought together computational-science teams, computer scientists, and the hardware vendor Intel. It was designated as one of two DOE high-performance computing research centers in 1993.

A second major achievement was the SNS. The lab had been designing a reactor-based neutron source. When the cost became unacceptable, Al took the lead in shifting to the accelerator-based SNS and gathered a multilab consortium to execute design and construction and broaden the project's base of support. The SNS returned the US to a world-leading position in neutron-scattering research.

Al was an advocate for international cooperation in science. His work behind the scenes with Evgeny Velikhov led to an agreement between the US and the Soviet Union on developing a magnetic-fusion demonstration model, which laid the groundwork for the ITER project.

AIP ESVA/PHYSICS TODAY COLLECTION

Alvin William Trivelpiece

Al's outstanding achievements extend across research, education, science administration, and national security. He was an innovator in the positions he held. Many of his accomplishments resulted from his understanding of science and politics and his ability to guide projects through government processes.

Those of us who worked closely with Al have been fortunate to have had the opportunity. Al was a continuous source of ideas and wisdom. He had a great sense of humor and was quick to generate puns. He enjoyed playing poker, chess, and golf and was an avid runner. He piloted his own airplane to the age of 85. His wisdom, leadership, and friendship will be missed.

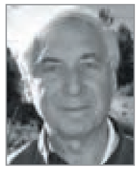
Thom Mason
Los Alamos National Laboratory
Los Alamos, New Mexico

James Decker
DGS & Associates LLC
Alexandria, Virginia PT

TO NOTIFY THE COMMUNITY about a colleague's death, visit https://contact.physicstoday.org and send us a remembrance to post.

Select submissions and, space permitting, a list of recent postings will appear in print.

QUICK STUDY



Jay Pasachoff (1943–2022) was the Field Memorial Professor of Astronomy at Williams College in Williamstown, Massachusetts. **Dan Seaton** is a principal scientist in the department of solar and heliospheric physics at the Southwest Research Institute in Boulder, Colorado. **Kevin Reardon** is a scientist at the National Solar Observatory and an adjunct professor at the University of Colorado Boulder.

Sunspots and their cycle

Jay M. Pasachoff, Daniel B. Seaton, and Kevin P. Reardon

These dark blemishes on stars are key to understanding the interaction of strong magnetic fields and dynamic plasma.

Looking through clouds or haze, people have glimpsed sunspots for thousands of years. But it was only with the magnification provided by the telescopes of Galileo Galilei and others that sunspots could be studied in detail and tracked as they crossed the solar disk. With that newfound ease and visibility, the science advanced quickly. As early as 1613, Galileo published a book with a wonderfully detailed sequence of drawings that showed the Sun's rotation.

In the mid 19th century, Heinrich Schwabe realized that the number of sunspots varies cyclically. Rudolf Wolf soon found that they peak roughly every 11 years, as shown in figure 1. In 1904, husband and wife team E. Walter Maunder and Annie Russell Maunder noticed that the distribution of spots also varies over time with latitude, and the plot they made, showing the decreasing latitude of spots during the solar cycle, resembled a butterfly with outstretched wings.

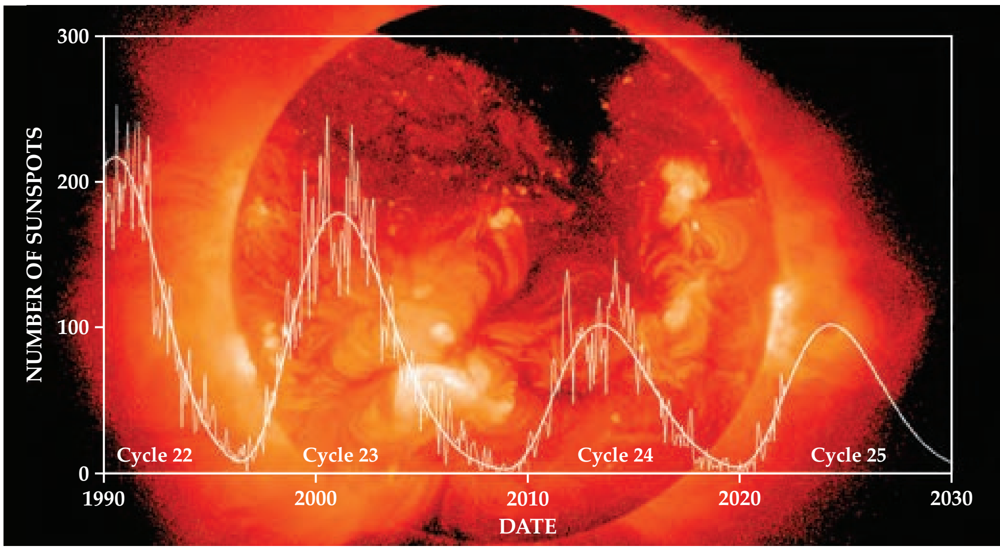


FIGURE 1. THE SUNSPOT CYCLE. The periodic change in the Sun's magnetic activity is measured by the variation in the number of sunspots on its surface. The rightmost part of the curve shows a prediction of the future sunspot activity. The background is an x-ray image from the Japanese *Yohkoh* spacecraft. At that wavelength, several-million-kelvin gas (orange) appears bright against regions of lower-density gas (black) extending toward the north pole—a magnetically defined "coronal hole," which is the source of the fast solar-wind streams. (Courtesy of David Hathaway and Lisa Upton.)

A measure of magnetism

In 1908, George Ellery Hale used a new spectrograph and solar telescope at the Mount Wilson Observatory to determine that sunspots are fundamentally a magnetic phenomenon. The strength of the field in the spots can be measured using the Zeeman effect, a splitting of certain spectral lines by an amount dependent on the amplitude of the magnetic field. Sunspots typically have magnetic field strengths of 1000–3000 gauss, thousands of times stronger than the Sun's global magnetic field.

The strong field reduces the energy brought to the surface by convection and makes the sunspot about 1000 K cooler than the surrounding photosphere—the Sun's visible surface (see figure 2). Though it would be brighter than the full Moon if it could shine against a dark sky, the sunspot region appears dark against the photosphere. Because of their reduced luminosity, the number of sunspots is one of the primary sources of variability in the total irradiance emitted by the Sun.

The National Solar Observatory's new Daniel K. Inouye Solar Telescope, with a 4-m aperture, sits on the rim of Haleakala Crater in Hawaii at an altitude of 3000 m. Its enhanced performance at IR wavelengths allows better studies of sunspots, given the increased sensitivity to magnetic fields of the Zeeman effect in this part of the spectrum.

Normally, sunspots occur in groups, and individual spots are linked to others of opposite magnetic polarity, with additional magnetic complexity being produced in part as magnetic field lines in the solar convection zone get twisted and kinked as they rise into the photosphere. Scientists actually track the solar activity using the Wolf sunspot number, the number of individual spots added to 10 times the number of different groups.

The daily sunspot number is recorded by several institutions, including the World Data Center at the Royal Observatory of Belgium, in Brussels. Its website, which displays the sunspot cycle traced back more than 300 years, is updated

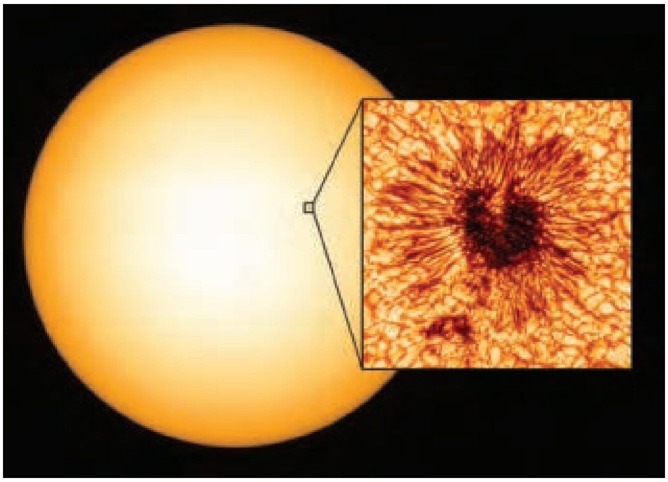


FIGURE 2. A SUNSPOT, UP CLOSE. This high-resolution image (inset) was taken at 530 nm by the Daniel K. Inouye Solar Telescope on 28 January 2020. The wavelength corresponds to the green-yellow part of the visible spectrum—close to the Sun's maximum brightness and, not coincidentally, the human eye's maximum sensitivity. The sunspot is produced by the joint action of intense magnetic fields and rising hot gas. Notice the dark umbra and surrounding fibrils (the penumbra), around which appear background solar granules. A wider view from the Global Oscillation Network Group telescopes shows the whole Sun and the relative size of the 15 000-km-wide (Earth-sized) sunspot. (Courtesy of NSO/NSF/AURA.)

monthly at http://sidc.oma.be/silso. Other sites tracking the sunspot cycle include the National Solar Observatory's Synoptic Optical Long-Term Investigations of the Sun (SOLIS), recently relocated to the grounds of the Big Bear Solar Observatory in California. SOLIS's suite of instruments provides a complete view of solar phenomena on a range of spatial scales—from hundreds of kilometers to the full solar disk—and on time scales from seconds to decades.

The sunspot cycle is actually a manifestation of a more general solar cycle of magnetic-linked activity, including solar prominences—relatively cool (a few 10 000 K) chromospheric plasma, suspended in the corona by the Sun's magnetic field—and coronal holes, regions of highly extended magnetic field lines, which are the source of the fast solar wind. The activity also results in the occurrence of abrupt solar flares, during which the coronal plasma might reach tens of millions of kelvin or be propelled into the solar system as coronal mass ejections.

Observatories and tracking

Every day at http://solarmonitor.org, you can see a variety of telescopic observations and links to other solar-activity data. The website includes magnetic field maps and extreme-UV images from NASA's *Solar Dynamics Observatory*. X-ray images from the defunct *Yohkoh* spacecraft and its current *Hinode* successor, for instance, show enhanced emission of hot (several million K) plasma in regions that overlay the sunspots, revealing the effects of concentrated magnetic fields in the corona.

In-depth studies of the evolution of the Sun's magnetic field reveal that the full magnetic solar cycle lasts 22 years. The Sun's differential rotation—faster at its equator than its poles—causes the field complexity to increase over a period of years. The increased complexity generates sunspots and coronal mass ejections, which release stored magnetic energy and allow the Sun's global magnetic field to return to its initial state, but with the magnetic polarity reversed, after about 11 years. (See the article by Gordon Holman, Physics Today, April 2012, page 56.)

In a given 11-year period between minima of solar activity, sunspot groups show a characteristic magnetic pattern of positive and negative polarities, with one polarity dominating in the leading spots in the group and the opposite polarity in the following spots (with an inverted sign in the northern and southern hemispheres). In the next 11-year cycle, that characteristic magnetic polarity behavior reverses until the Sun returns again to its initial state after more than two decades.

Over the past three decades, a system of solar observing stations known as the Global Oscillation Network Group has been mapping the Sun's surface from six sites longitudinally distributed around Earth and provides images nearly continuously. Space-based observatories, such as NASA's *Parker Solar Probe* and the European Space Agency's *Solar Orbiter*, will continue to orbit the Sun ever more closely over the next few years.

Solar Orbiter has captured the first images of sunspots in the photosphere from a new perspective, away from the Sun–Earth line. In 2024, *Parker Solar Probe* will pass only 6.2 million km above the Sun's surface, making it the first space mission to sample plasma and magnetic fields in the Sun's extended atmosphere directly (see the article by Nour Raouafi, Physics Today, November 2022, page 28).

In addition to the new Inouye Solar Telescope, other terrestrial observatories with capabilities to take high-resolution sunspot images include the 1-m Swedish Solar Telescope and the Leibniz Institute for Solar Physics's 1.6 m Gregor telescope, both in the Canary Islands, and the New Jersey Institute of Technology's 1.6 m Goode Solar Telescope at the Big Bear Solar Observatory. Those telescopes all use adaptive optics to correct atmospheric distortions and sharpen the sunspot images.

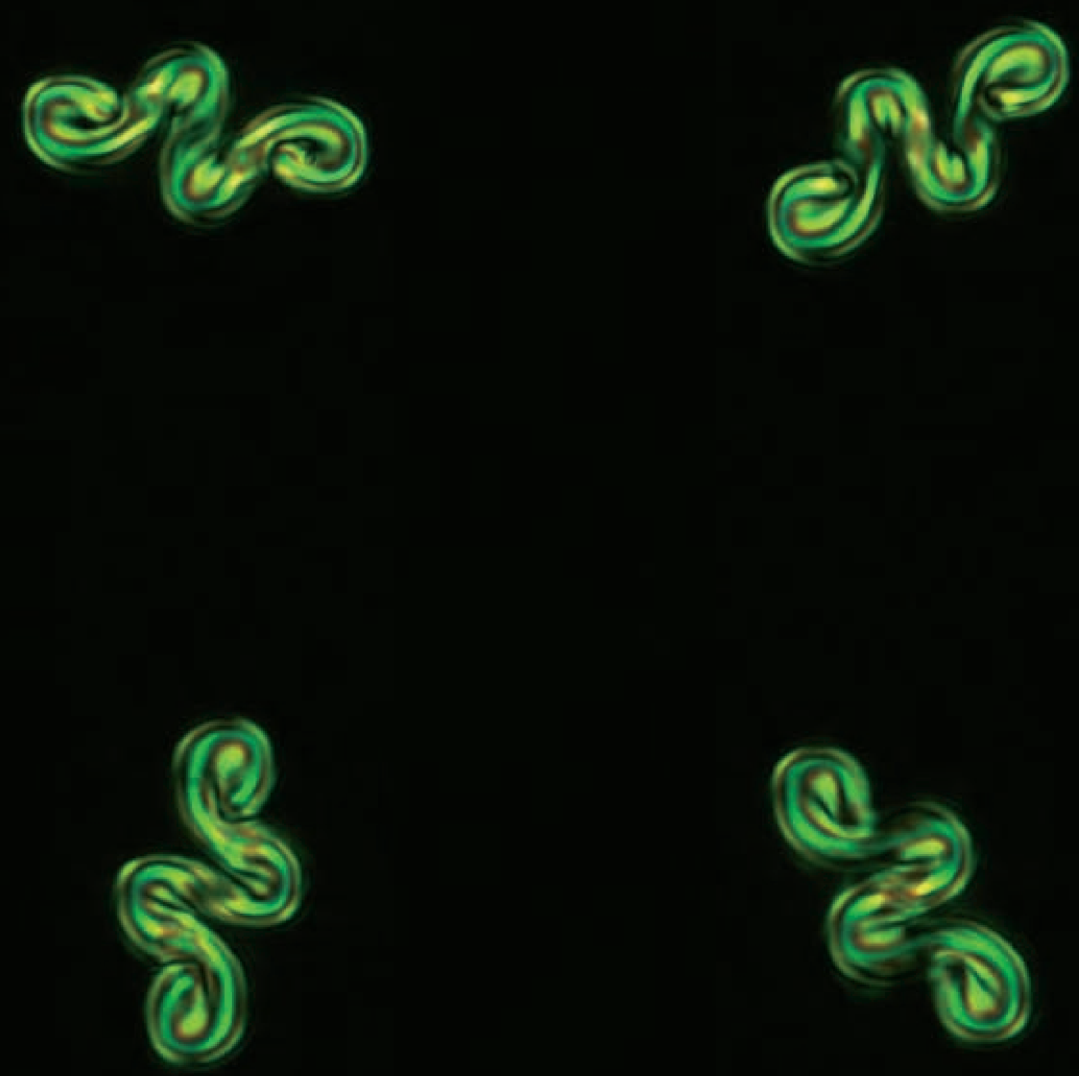
But even casual observers—using simple, safely filtered solar telescopes—can, like Galileo, follow sunspots day-to-day or year-to-year and get a sense for the complicated magnetic processes happening in the solar atmosphere.

Additional resources

- ▶ National Center for Atmospheric Research High Altitude Observatory, https://www2.hao.ucar.edu/.
- ▶ Solar Physics Group at NASA's Marshall Space Flight Center, https://solarscience.msfc.nasa.gov.
- ▶ The European Space Agency's *Solar Orbiter*, https://www.esa.int/Science_Exploration/Space_Science/Solar_Orbiter.
- ▶ Daniel K. Inouye Solar Telescope, https://nso.edu/telescopes/dki-solar-telescope/.
- ▶ J. Brody, *The Enigma of Sunspots: A Story of Discovery and Scientific Revolution*, Floris Books (2002).
- ▶ L. Golub, J. M. Pasachoff, *The Sun*, U. Chicago Press (2017).
- ▶ J. B. Zirker, *Sunquakes: Probing the Interior of the Sun*, Johns Hopkins U. Press (2003).

PT

BACK SCATTER

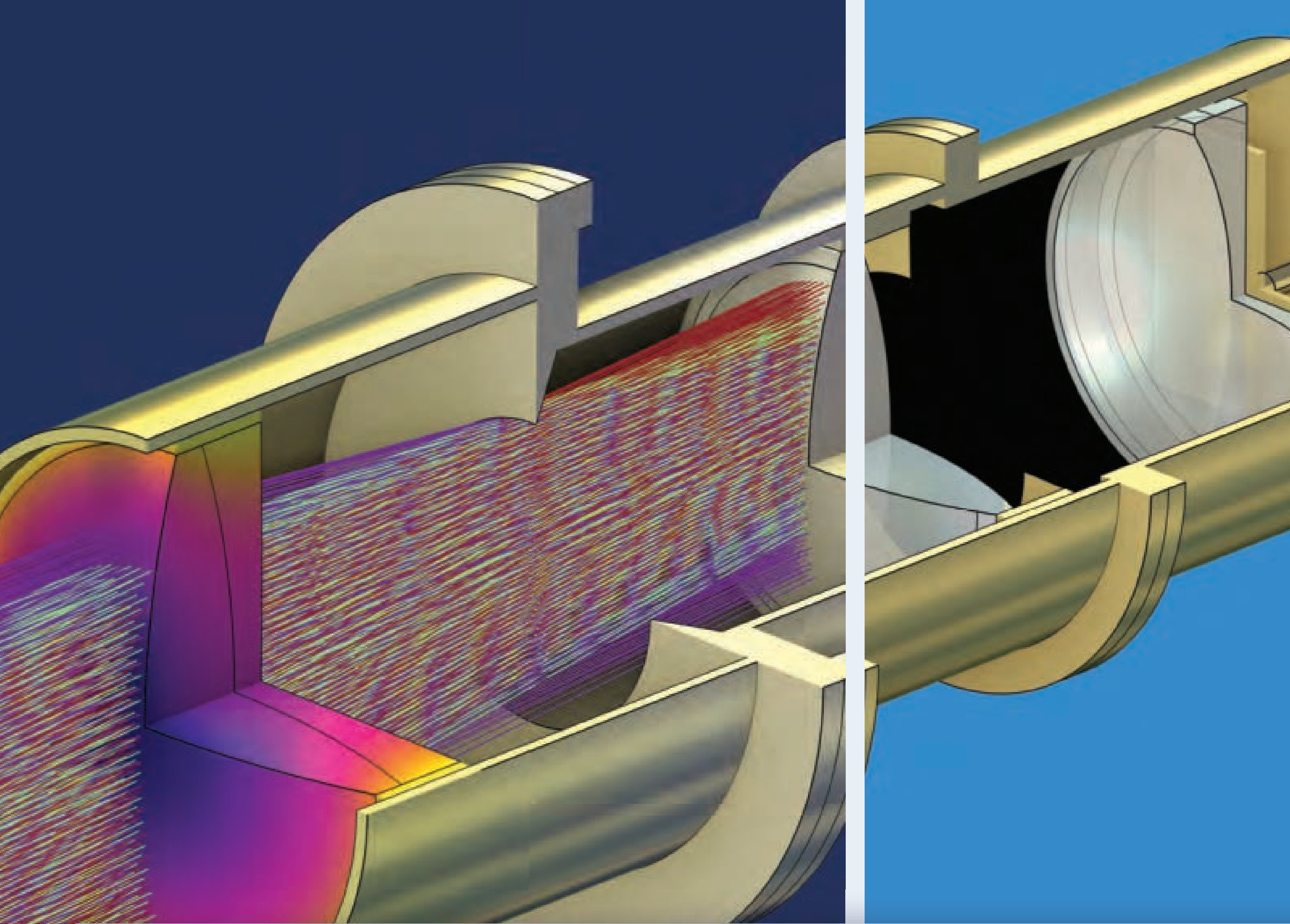


Structures like Möbius strips

A Möbius strip defies orientation. If a two-dimensional asymmetric object were to travel around such a strip, it would return to the starting point but as a mirror image of itself. During the course of their topological soft-matter research, PhD student Hanqing Zhao and his adviser Ivan Smalyukh of the University of Colorado Boulder generated 3D structural defects reminiscent of a Möbius strip. Shown here are four of what the researchers call möbiusons, each about 10 μm long. They form in certain nonpolar, chiral liquid crystals when line defects, also known as vortices, self-assemble with topological solitons, the continuous localized twists that form when molecules align.

Liquid crystals are composed of rod-shaped molecules that align in an orientational order—most of the molecules point in the same direction. But möbiusons have nonorientable structures that can spontaneously fold into a diverse set of spatially localized configurations. When Zhao and his colleagues applied an electric field, they observed self-propelled möbiusons that were stable through various rotational and translational motions. Such behavior could potentially be harnessed to transport nano-sized cargo. The researchers suspect that the topology-based design of their möbiusons could also organize nanoparticles into mesoscale spatial patterns. (H. Zhao et al., *Nat. Phys.*, 2023, doi:10.1038/s41567-022-01851-1; images courtesy of Hanqing Zhao.) —AL

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



MathWorks®

MATLAB SPEAKS DEEP LEARNING

With MATLAB®, you can build deep learning models using classification and regression on signal, image, and text data. Interactively label data, design and train models, manage your experiments, and share your results.

mathworks.com/deeplearning

Semantic segmentation for wildlife conservation.