PHYSICS TODAY

June 2020 • volume 73, number 6

A publication of the American Institute of Physics

ATOMIC CLOCKS for GRAVITATIONAL REDSHIFTS

Majorana quasiparticles

Protein-based logic gates

Electron hydrodynamics

APSIT
INSURANCE FOR SCIENCE PROFESSIONALS

What old-fashioned ideas are you still holding on to?

If you still think life insurance is too expensive, take another look at APSIT Group 10-Year Level Term Life Insurance.

DID YOU KNOW?

Premiums can start at $9.00 a month[1] and don't fluctuate, even if your health changes. That's ten years of protection for your loved ones around the price of a cassette tape in 1985!

APPLY ONLINE TODAY!

APSITPLANS.COM/LTL-NOW | 800.272.1637

[1] The Preferred rate shown is calculated based on $250,000 of 10-year level term life coverage for a non-smoking, healthy female, 30-35 years of age. A $0.50 administrative fee will be applicable to all premium payment modes other than annual.

Underwritten by New York Life Insurance Company, 51 Madison Avenue, New York, NY 10010 on policy form GMR. For complete details on APSIT 10-Year Level Term Life Insurance, including features, costs, eligibility, renewability, limitations, and exclusions, see the Certificate of Insurance.

Program Administrators: Arkansas Insurance License #1322, California Insurance License #0F76076

192136-APSIT-LTL-MAG-PAD


PHYSICS TODAY
OCTOBER 2020
MARK YOUR CALENDAR
2ND ANNUAL CAREERS & RECRUITMENT ISSUE

PHYSICS TODAY

June 2020 | volume 73 number 6

FEATURES

ON THE COVER: With a height of 634 meters, the Tokyo Skytree broadcasting tower is the tallest structure in Japan. A team of researchers developing transportable atomic clocks used that height to put their clocks to the test. By placing one clock at the tower's base and the other at the observatory level, they precisely measured the relativistic frequency shift between the clocks. For more about the transportable clocks, see the story on **page 20**. (Sean Pavone/Alamy Stock Photo.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org

ISTOCK.COM/TOMBIE

Reality check
Mathematical reasoning leads to some befuddling possibilities, including the assembly of a Sun-sized sphere from a pea-sized one or a monkey churning out Hamlet on a typewriter. Alas, the real world gets in the way. Johanna Miller explains why in her latest column.
physicstoday.org/Jun2020a

COURTESY OF DAISY SHEARER

Autism ambassador
Last year physicist Daisy Shearer publicly shared her autism diagnosis on Instagram and began using social media to mentor and advocate for neurodivergent people in STEM fields. Shearer talks to Physics Today about her experiences, her spintronics research, and her efforts to encourage inclusion.
physicstoday.org/Jun2020b

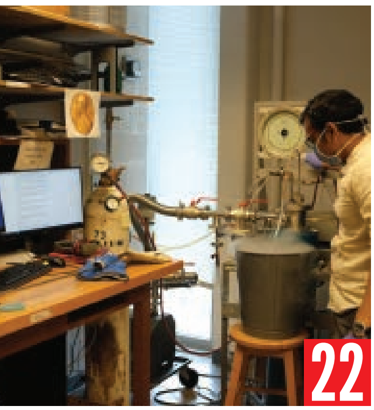
ESO

Distributed review
Under a policy called distributed peer review, applicants for research grants or telescope time evaluate each other's proposals. A recent trial found that reviewers judging fellow applicants gave proposal critiques similar to those by an independent group of peer reviewers.
physicstoday.org/Jun2020c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 73, number 6. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 300, 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 300, 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**

Senior director of news & magazines
Larry Fishbein *lfishbein@aip.org*

Editor-in-chief
Charles Day *cday@aip.org*

Managing editor
Richard J. Fitzgerald *rjf@aip.org*

Art and production
Donna Padian, art director
Freddie A. Pagani, graphic designer
Cynthia B. Cummings, photographer
Nathan Cromer

Editors
Melinda Baldwin *mbaldwin@aip.org*
Toni Feder *tf@aip.org*
Martha M. Hanna *mmh@aip.org*
Heather M. Hill *hhill@aip.org*
David Kramer *dk@aip.org*
Alex Lopatka *alopatka@aip.org*
Christine Middleton *cmiddleton@aip.org*
Johanna L. Miller *jlm@aip.org*
Gayle G. Parraway *ggp@aip.org*
R. Mark Wilson *rmw@aip.org*

Online
Paul K. Guinnessy, director *pkg@aip.org*
Andrew Grant, editor *agrant@aip.org*
Angela Dombroski *atd@aip.org*
Greg Stasiewicz *gls@aip.org*

Assistant editor
Cynthia B. Cummings

Editorial assistant
Tonya Gary

Contributing editor
Andreas Mandelis

Sales and marketing
Christina Unger Ramos *cunger@aip.org*
Unique Carter
Krystal Dell
Skye Haynes

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

PhysicsToday **@physicstoday**

AIP | American Institute of Physics

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

Other member organizations
Sigma Pi Sigma Physics Honor Society
Society of Physics Students
Corporate Associates

FROM THE EDITOR

In praise of theorists who build bridges

Charles Day

For PHYSICS TODAY's August 2006 issue I wrote a news story about the observation of the Berezinskii-Kosterlitz-Thouless topological phase transition in flattened clouds of ultracold rubidium atoms. As conceived in the early 1970s by Vadim Berezinskii, J. Michael Kosterlitz, and David Thouless, the BKT transition occurs in a two-dimensional lattice model known as XY. When I wrote the story, the transition had already been observed in superfluid films of helium-4 and superconducting films of mercury–xenon alloy.

The 2006 experiment was conducted at the École Normale Supérieure (ENS) in Paris by Zoran Hadzibabic, Peter Krüger, Marc Cheneau, Baptiste Battelier, and Jean Dalibard. In summarizing its significance, I wrote, "Their experiment not only confirms BKT theory in a new system, but also reveals for the first time the transition's microscopic instigators: local topological defects or vortices."

Then as now, the editors who write for PHYSICS TODAY's Search and Discovery department choose papers to cover based primarily on the advice of experts. We don't learn a paper's backstory until we interview its authors. When I phoned Hadzibabic to ask about his experiment, I found out something unexpected and interesting. To connect the predictions of BKT to the measurements made in the lab, he and his collaborators relied on a method devised by an independent trio of theorists, Anatoli Polkovnikov, Ehud Altman, and Eugene Demler.[1] As in the plot of a B movie, a preprint from the three theorists arrived at ENS just as the experimenters were wondering how to interpret their data. "Within a minute, we wrote back to say we have the same equation on the board!" Hadzibabic told me.

Polkovnikov, Altman, and Demler took up the challenge after reading a previous paper by the ENS group. The correlations that embody the BKT transition are manifest when two pancakes of ultracold atoms are released from their traps and allowed to interfere with each other. Those correlations, Polkovnikov, Altman, and Demler realized, depend on system size in an experimentally accessible way that can be tied directly to BKT physics.

One of my favorite examples of theorists building bridges between experiment and theory comes from solar physics. Our star's photosphere is so dense that a photon emitted at the core takes 170 000 years of repeated scattering off electrons and ions before it escapes. Some of the sunlight that fell on your face today was born when Neanderthals hunted and gathered in Europe! If astronomers relied only on solar photons, they'd have little to test their theories of the Sun's composition and structure. Fortunately, prompt, direct information from the Sun's interior reaches us in two other forms: acoustic oscillations, which I discuss here, and solar neutrinos.

The Sun is a self-gravitating, differentially rotating ball of plasma that quivers in myriad acoustic modes, some of which entrain matter deep in the solar interior. Those helioseismic signals are manifest as Doppler shifts of certain spectral lines at localized patches in the photosphere. Remarkably, those seemingly limited data are sufficient to constrain models of the Sun—provided someone goes to the trouble of collating, among other things, all the nuclear reactions of all the chemical elements at all levels of the Sun and then creating models that predict what can be detected.

The effort of helioseismologists is prodigious. To give one example, in a 1991 paper, Jørgen Christensen-Dalsgaard, Douglas Gough, and Michael Thompson derived the depth of the Sun's convection zone.[2] To get to their answer of 0.287 ± 0.003 solar radii, they started with the Schwarzschild criterion, which specifies when convention ensues. Over the course of 25 pages, the three theorists described how they built two models of the Sun and then computed the sound speed, which can be inferred from helioseismic observations.

I mention the work of those bridge-building theorists not just to praise them and their kind. When I look back at my own physics education, I don't recall being taught what might be called applied theory. Polkovnikov has recognized the same gap. When I sought his comments about this editorial, he told me that he, Marcos Rigol, and Pieter Claeys are writing a new quantum mechanics textbook that will include realistic, nonideal examples. "When the formalism is too abstract and is related to very particular experiments," he wrote, "it creates a gap in intuition and a gap in connecting theory and experiment."

References

1. A. Polkovnikov, E. Altman, E. Demler, *Proc. Natl. Acad. Sci. USA* **103**, 6125 (2006).
2. J. Christensen-Dalsgaard, D. O. Gough, M. J. Thompson, *Astrophys. J.* **378**, 413 (1991).

PT

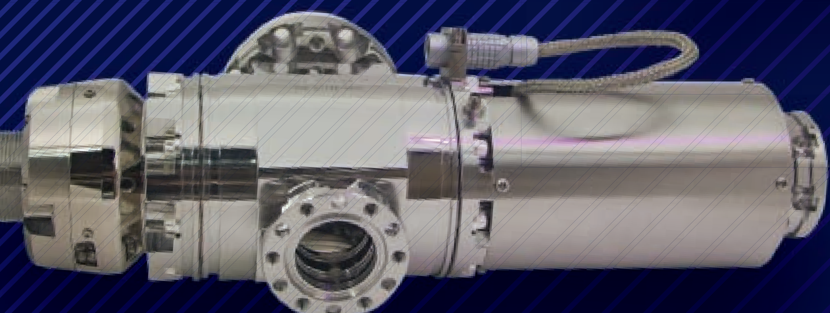
KIMBALL PHYSICS

Excellence in Electron and Ion Optics

EFG - 7 (1500 eV)

ELECTRON GUN SYSTEMS

Flood beams, small spot
High brightness LaB_6
Ultra-fast ps to fs Photo Emission Guns
1 eV to 100 keV



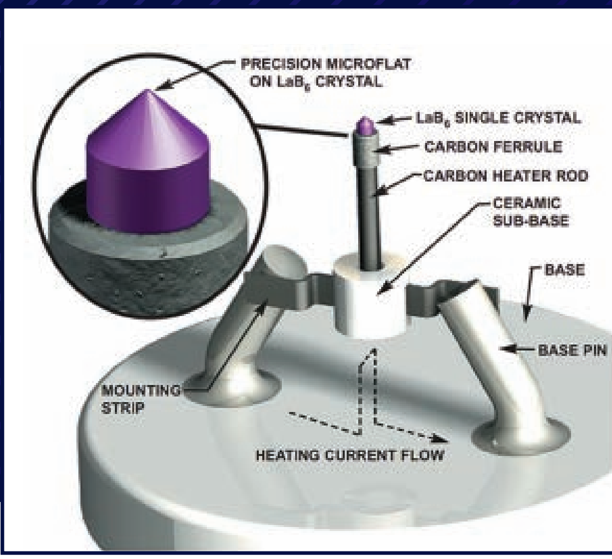
EGH - 6210 (60 keV)

- LaB_6 for FEL experimental brain surgery
- MBE gun for Space Shuttle missions
- CubeSat low power emitters

- Electron guns for solar wind studies on satellite components
- Custom chambers
- Magnetic optical column for semiconductor inspection and high energy Transmission Electron Microscope

Stainless Steel

Titanium



UHV MULTI -CF FITTINGS
Some of the over 100 Available
Custom Encouraged
Metric and Titanium Available

www.kimballphysics.com
311 Kimball Hill Road, Wilton,NH 03086-9742 USA

Kimball Physics is now celebrating its 50th Anniversary!

Changing the paradigm for research publishing

ISTOCK.COM/COFOTOISME

TODAY'S RESEARCH PUBLICATION MODEL was developed when journals appeared on paper only and submissions were either accepted or rejected. Perhaps it's time to consider a new model.

The commentary by Detlef Lohse and Eckart Meiburg, "On the quality and costs of science publication" (PHYSICS TODAY, August 2019, page 10), criticizes the Plan S initiative of some European funding agencies, which would require that results of publicly funded research be published in open-access journals. Lohse and Meiburg's points are fair, given the present research publication paradigm.

The point they are missing, to my mind, is that the paradigm is—and should be—changing. Many aspects of the present model come from the time when journals appeared only on physical paper, and some ways are so deeply rooted in the community that they are hardly questioned. Among them are the enormous and ever-growing number of journals and our reliance on the use of journal names to screen for quality. I think those practices are neither optimal nor indisputable.

Instead of questioning Plan S based on the existing publishing model, I see the plan as an opportunity to revise the model. For example, a key criticism to open-access publishing is that it favors bad journals. That is, publishing in "good" journals that are highly selective is expensive, whereas publishing in nonselective "bad" journals is much less costly. That is true if we do not question the present accept-or-reject publishing model. However, if we consider using peer review for quality discrimination by grading papers instead of rejecting them, the scenario changes radically. Imagine an all-physics journal like the new *Physical Review Research* using grades for its papers to correlate with levels in the old model—including letters (grade 3), rapid communications (grade 2), regular articles (grade 1), and even higher and lower levels. Not only would such a journal be perfectly geared for Plan S, since the rejection rate would be minimized and publication costs thereby reduced, but it would also represent significant progress in research publishing, as explained in www.emilio-artacho.blogspot.com.

Emilio Artacho
(ea245@cam.ac.uk)
University of Cambridge
Cambridge, UK

Getting the drop on quantum droplets

The article "Ultradilute quantum droplets" by Igor Ferrier-Barbut (PHYSICS TODAY, April 2019, page 46) was really nice to see. It reported on the creation, at last, of real ultradilute liquid droplets and on the tremendous progress that has been made in that area. However, I was disappointed to see no mention or discussion that ultradilute liquid quantum droplets were predicted a long time ago.[1,2] That early work was a major source of inspiration for Dmitry Petrov's 2015 paper on the subject,[3] at least according to what he told me years ago.

In his discussion of mean-field quantum gases, Ferrier-Barbut doesn't note that the Efimov effect, which involves the creation of an infinite number of three-body bound states, can also allow Bose liquids to exist when the scattering length is negative (that is, when the two-particle interaction is attractive). Such a system is not always unstable. Tsung-Dao Lee, Kerson Huang, and Chen Ning

Yang (LHY) proposed in 1957 a leading-order correction to the mean-field approximation due to two-body collisions, which was used by Petrov;[3] that "game-changing correction," as Ferrier-Barbut calls it, can alter the nature of the Bose–Einstein condensate.

However, as was discussed quite some time ago,[1,2] the strength of three-body interactions can dominate over LHY corrections and can be infinite even if the two-body scattering length is finite.[1] A liquid model based on the Efimov effect[1,2] is more robust than the one Petrov envisioned and much more flexible than the van der Waals model. Unlike the quantum liquid droplets created in mixtures of Bose–Einstein condensates,[4] which have practically the same size for particle numbers up to tens of thousands, the quantum liquid droplets I suggested are truly saturating systems, with basically constant interior density. A droplet can have any size, and it can be formed even from a single element. It is a real liquid, with constant density inside and a well-defined surface, and its density and surface tension can be controlled. Also, it is stable against quantum corrections to the mean field.[2]

Moreover, in a rather special system—an ensemble of spin-polarized tritium atoms—three-body recombination processes are most likely absent.[2,5] Although I did not make the estimates, which should be straightforward, I am sure that by controlling the density and thus the rate of four-body recombination, one could create droplets with basically arbitrarily long lifetimes. A droplet of spin-polarized tritium atoms would be a totally unique object, perhaps as unique as macroscopic superfluid helium, but amenable to precise quantum many-body calculations, both static and time-dependent. Quantum turbulence could be studied in a large class of systems, for which a microscopic theory exists, and unlike in the case of superfluid helium, theory could be directly confronted with experiment.

Quantum liquid droplets could be either boselets or fermilets and would undergo at least two types of phase transitions, from superfluid to normal and from liquid to gas. Their physics should be fascinating. Mixing bosons and fermions can lead to even more interesting and complex objects.

References

1. A. Bulgac, *Phys. Rev. Lett.* **89**, 050402 (2002).
2. P. F. Bedaque, A. Bulgac, G. Rupak, *Phys. Rev. A.* **68**, 033606 (2003).
3. D. S. Petrov, *Phys. Rev. Lett.* **115**, 155302 (2015).
4. C. R. Cabrera et al., *Science* **359**, 301 (2018).
5. D. Blume et al., *Phys. Rev. Lett.* **89**, 163402 (2002).

Aurel Bulgac
(bulgac@uw.edu)
University of Washington
Seattle

▶ **Ferrier-Barbut replies:** I am grateful for Aurel Bulgac's insight about three-body stabilized quantum droplets. I was aware of his work, but space constraints made it impossible for me to cite the broad swath of related literature. A tritium droplet would certainly be a peculiar object, though as an experimentalist I think making a Bose–Einstein condensate of tritium would be quite challenging.

Igor Ferrier-Barbut
(igor.ferrier-barbut@institutoptique.fr)
Institut d'Optique
CNRS
Palaiseau, France

Reviews of quantum foundations

I enjoyed the February 2019 issue of PHYSICS TODAY on *Reviews of Modern Physics* at 90 but was disappointed with the article "Quantum foundations" by David DiVincenzo and Christopher Fuchs (page 50). The most useful part of that article was the reference list, which shows *RMP*'s diversity of papers on the subject. My 1970 article on the statistical-ensemble interpretation of quantum mechanics (QM),[1] which people tell me has encouraged them to continue research on quantum foundations (QF), was omitted from the list.

Unfortunately, DiVincenzo and Fuchs continue to mystify measurement in QM, as if it were some deep philosophical concept that must be treated before QM has even been fully formulated. They assert that "physicists and philosophers are still debating what a 'measurement' really means." What is important for QF is not the meaning of the word but an understanding of the physical process. The authors do not cite any of the published papers that provide such an understanding. And they give too much attention to two marginal interpretations: the many-worlds interpretation (MWI) and quantum Bayesianism (QBism).

In QM, a measurement of an observable should yield an eigenvalue of the observable. If the initial state of the measured object is a superposition of eigenstates corresponding to different eigenvalues, then the interaction of the measurement apparatus with the object will lead to a final state of the whole system—measured object plus apparatus—that is a superposition of different measurement results. The squared amplitude of each term yields the probability of obtaining that result in an individual measurement. That statistical prediction, the Born rule, is common to the Copenhagen and statistical-ensemble interpretations. But the MWI takes a radically different turn. It postulates that the universe branches into several parallel worlds, with each term of the superposition corresponding to the unique result of the measurement in one branch world.

The usual role of an interpretation of QM is to begin with the established mathematical formalism and provide an intuitively comprehensible idea of the physical process that the math describes. The MWI does not do that. Instead, it adds a mysterious process of world-splitting, a strange new cosmology that is alien to the mathematics of QM and not really an interpretation of QM at all. A typical QM measurement, such as that of a spin component in the Stern–Gerlach experiment, is a local and very low energy event. It is not credible that the measurement could have the huge cosmological effect of bifurcating the universe.

When I first heard of the world-splitting assumed in the MWI, I went back to Hugh Everett's paper[2] to see if he had really said anything so absurd. I found that he had not said so explicitly, but he sometimes used words that could be interpreted in more than one way. The MWI is a possible interpretation of them, but not the most natural one, so I thought. And Everett's framework still has value even without resorting to the MWI's world-splitting. His concept of a "relative state" is useful, for instance, and he is correct in rejecting the notion of the quantum state "collapsing" after a measurement.

QBism begins with the assumption that all kinds of probability can be regarded as subjective Bayesian probabilities. That assumption can be maintained only by ignoring the literature on interpretations of probability, from which it is clear that several different kinds—or interpretations—of probability exist. DiVincenzo and Fuchs may have ignored the classic philosophical writings on the subject because they were written by philosophers for philosophers and so do not address the needs of physicists.

I have published a paper on the foundations of probability theory, written from the point of view of a quantum physicist.[3] I classify the main kinds or interpretations of probability into three groups: inferential probability, of which Bayesian theory is an example; frequency or ensemble probability, commonly used in Gibbsian statistical mechanics and in QM; and propensity theory. Propensity, a degree of causality that is weaker than determinism, is not merely another interpretation of probability. Its mathematical theory must also differ from that of probability theory, as Paul Humphreys showed[4] in 1985. Although the axioms of propensity[3] differ from those of probability, the two axiom sets overlap. Both support the law of large numbers, so propensity theory is compatible with the most useful part of the frequency interpretation of probability.

In general, QM states do not determine the results of a measurement, only the probabilities of the possible results. That a state's influence on the results is not deterministic suggests strongly that the quantum probabilities given by the Born rule should be interpreted as propensities. They refer objectively to the physical system and its environment, not to any agent's knowledge, so they are not naturally interpreted as subjective Bayesian probabilities.

Interpretations of probability may differ not only in philosophy but also in substance. As I discuss in reference 3, John Bell's theorem illustrates how local hidden-variable theories are incompatible with QM. E. T. Jaynes was a well-known supporter of the Bayesian theory of probability. In 1989 he repeated Bell's derivation of inequality but carefully treated all instances of probability as Bayesian. He found that the derivation could not be completed without invoking an extra assumption that was not justifiable in the Bayesian theory. Bell's theorem involves questions about causality, so it is natural to use propensity theory to treat it. That method is successful in deriving Bell's inequality.[3]

Not all probabilities occurring in QM can be treated as subjective Bayesian probabilities. That limitation disqualifies QBism, a Bayesian-based theory, as an interpretation of QM that can succeed in quantum foundations. The initial assumption of QBism is not valid.

References

1. L. E. Ballentine, *Rev. Mod. Phys.* **42**, 358 (1970).
2. H. Everett III, *Rev. Mod. Phys.* **29**, 454 (1957).
3. L. E. Ballentine, *Found. Phys.* **46**, 973 (2016).
4. P. Humphreys, *Philos. Rev.* **94**, 557 (1985).

Leslie Ballentine
(leslie_ballentine@sfu.ca)
Simon Fraser University
Burnaby, British Columbia, Canada

▶ **DiVincenzo and Fuchs reply:** We deeply regret our oversight of Leslie Ballentine's influential 1970 *Reviews of Modern Physics* article on the ensemble interpretation. We were well aware of the paper but had not realized that it appeared in *RMP* so as to be appropriate for the retrospective. We apologize to Ballentine and to our readership.

David P. DiVincenzo
(d.divincenzo@fz-juelich.de)
Peter Grünberg Institute
Jülich, Germany

Christopher A. Fuchs
(christopher.fuchs@umb.edu)
University of Massachusetts Boston

Celestial background of 1869 eclipse

I enjoyed Deborah Kent's article on American efforts to document and study the 1869 total solar eclipse (Physics Today, August 2019, page 46). At the April 2019 meeting of the American Physical Society, we were treated to a session titled "Centennial of the Eddington Eclipse Expedition."

I'm curious. Were stars visible in any of the photos of the 1869 eclipse—or other eclipses in the days before general relativity? And would it have been possible that someone noticed the displacement of the stars' positions as Arthur Eddington did in 1919, but before Albert Einstein published his theory in 1915?

Robert McAdory
(robertmcadory@yahoo.com)
Clinton, Mississippi

▶ **Kent replies:** I'm glad Robert McAdory enjoyed the article. Although I'm not aware of any photos from the 1869 eclipse that show visible stars, there were images on plates from eclipses before 1919. Expeditions from the Lick, Yerkes, Smithsonian Astrophysical, and US Naval Observatories took large-format images of the corona during the 28 May 1900 eclipse, when the star field was similar to that during the 29 May 1919 eclipse. The images taken by Lick and their possible connections to the relativity test are explored in chapter five of Jeffrey Crelinsten's *Einstein's Jury: The Race to Test Relativity* (2006) and chapter two of *No Shadow of a Doubt: The 1919 Eclipse That Confirmed Einstein's Theory of Relativity* (2019) by Daniel Kennefick.

The 19th-century searches for an intramercurial planet resulted in many images in which some background stars might be visible. The *Lick Observatory Bulletin*, number 24 (1902), reported that half of the observatory's plates from the eclipse of 18 May 1901 included star images. That report also has more specific information about Lick's capabilities to capture stars in images.

My thanks to Tom English of the Cline Observatory, Jamestown, North Carolina.

Deborah Kent
(deborah.kent@drake.edu)
Drake University
Des Moines, Iowa

Accelerator experiments are closing in on neutrino *CP* violation

It's starting to look like neutrinos and antineutrinos aren't exact mirror images of each other.

©J-PARC CENTER

FIGURE 1. WHERE NEUTRINOS ARE MADE. As charged pions created at the J-PARC accelerator in Japan fly through this 96-m-long tunnel, they decay into muons and muon neutrinos. The wall at the end, 5 m tall by 3 m wide, stops the muons and any undecayed pions. The neutrinos, which have no problem traveling through solid metal and rock, keep going toward the detector 295 km away.

Somewhere in the laws of physics, particles must be allowed to behave differently from their antiparticles. If they weren't, the universe would contain equal amounts of matter and antimatter, all the particles and antiparticles would promptly annihilate one another, and none of us would exist.

Violations of *CP* symmetry—the combination of charge conjugation and parity inversion that projects particles onto their antiparticles' mirror images—have already been observed and theoretically accounted for in several flavors of quarks. (See PHYSICS TODAY, August 2019, page 14.) But the extent of that violation is nowhere near enough to explain the imbalance of matter and antimatter in the universe. To make up the shortfall, researchers are looking for an additional source of *CP* violation among the particles of the lepton sector: electrons, muons, taus, all their antiparticles, and their associated neutrinos and antineutrinos.

Unlike most other known particles, neutrinos spontaneously change their identities as they travel. (See PHYSICS TODAY, December 2015, page 16.) A muon neutrino created in one place, for example, might later be detected as an electron neutrino in another. The dynamics of that flavor oscillation can be characterized by three mixing angles—θ_{12}, θ_{23}, and θ_{13}—plus a phase δ_{CP} that captures the amount of *CP* violation, if any. A value of 0 or $\pm\pi$ radians for δ_{CP} means that neutrinos and antineutrinos oscillate identically and *CP* symmetry is conserved; any other value means that the symmetry is broken.

Theory leaves the values of all four of those parameters wide open. And they're extremely difficult to measure experimentally, because neutrino oscillations are extremely difficult to detect. The three mixing angles have been measured to within a few degrees. But the value of δ_{CP} has remained almost entirely unknown.

Now the Tokai-to-Kamioka (T2K) experiment is homing in on δ_{CP}. By smashing protons into a graphite target at the J-PARC accelerator in Tokai on Japan's east coast, the researchers create a powerful, steady beam of either muon neutrinos or muon antineutrinos. At the Super-Kamiokande detector, 295 km to the west, they measure how many of those particles have changed flavor. By comparing the results from neutrino and antineutrino beams, they can estimate δ_{CP}.

After 10 years of data collection—interrupted, unfortunately, by the Tohoku earthquake and tsunami that devastated Japan in 2011 and by an accident on another beamline at J-PARC in 2013—the T2K data suggest that muon neutrinos transform into electron neutrinos more readily than muon antineutrinos transform into electron antineutrinos.[1] It's not yet the end of the story: Whereas some values of δ_{CP} are excluded at a confidence level of three standard deviations (3σ), perfect *CP* symmetry is disfavored only with 2σ confidence. A conclusive answer, by convention, requires 5σ confidence.

But the result suggests that experimental searches for lepton *CP* violation are probably on the right track and could reach that threshold in the coming years as data collection continues and new facilities come on line.

The matter of matter

Neutrinos don't often make their presence felt. Apart from the rare weak interaction, they stream—unseen and untouched—through space, through solid rock, and through the densest plasma of the Sun. Odds are good that none of the atoms in your body will ever interact with a neutrino during your lifetime. How can such an aloof bunch of particles have anything to do with the existence of all the matter in the universe?

A speculative answer lies in the neutrino masses. (See the Quick Study by Rabi Mohapatra, PHYSICS TODAY, April 2010, page 68, and the article by Helen Quinn, PHYSICS TODAY, February 2003, page 30.) Each neutrino flavor state is a quantum superposition of the same three mass states, which beat in and out of phase with one another and result in flavor oscillations. The masses' exact values are unknown, but they seem to be on the order of millielectron volts. Neutrinos are thus the outliers among massive fundamental particles, the rest of whose masses are best measured in megaelectron volts or gigaelectron volts.

Some theories postulate that whatever mechanism gives neutrinos their anomalously small masses also creates another set of particles with anomalously large ones. The hypothetical particles would be too massive ever to be seen today, even in the most powerful of particle accelerators, but they could have been abundant in the energetic environment of the early universe. If they existed, and if their decays into other particles violate *CP* symmetry, they could be responsible for the present-day matter–antimatter imbalance. And if ordinary neutrinos and their ultraheavy counterparts behave the same way with respect to *CP* symmetry, then studying neutrinos could open a window onto the dynamics of the early universe.

That's a lot of ifs, and not all theoretical models accommodate that line of reasoning. But for those that do, looking for neutrino *CP* violation is a place to start.

Neutrino beam

Much of what we know about neutrinos comes from experiments that monitor the neutrinos coming from nuclear reactors. (See, for example, PHYSICS TODAY, May 2012, page 13.) When the unstable fission fragments made in a nuclear reaction undergo beta decay, they spew out multitudes of electron antineutrinos in all directions. Accelerator experiments offer an additional degree of control: the ability to create an intense beam of neutrinos that all travel in approximately the same direction.

Neutrino creation at J-PARC is efficient. The accelerator protons colliding with the graphite target create a multitude of charged pions, almost all of which decay into muons and muon neutrinos as they pass through the tunnel shown in figure 1. The pions and muons propagate through the free space of the tunnel but are stopped by the solid wall at the end, leaving only the neutrinos to continue the trip west to Super-Kamiokande. A 40-m-tall water tank lined with photomultiplier tubes, Super-Kamiokande is one of the best neutrino detectors in the world.

When a neutrino passing through Super-Kamiokande chances to undergo a weak interaction with one of the water molecules, it produces a charged lepton—an electron or muon—according to the neutrino's flavor. (Taus take too much energy to produce, so tau neutrinos mostly go undetected.) From the Cherenkov light the lepton generates as it speeds through the water, the researchers can determine the neutrino's flavor and, crucially, exactly when it arrived. Super-Kamiokande got its start detecting neutrinos coming from the Sun and from Earth's atmosphere, and those neutrino sources are still there. J-PARC creates neutrinos in precisely timed pulses; based on their time of arrival, Super-Kamiokande distinguishes the neutrinos coming from the accelerator and those from all other sources.

Only a small fraction of the neutrinos from J-PARC get detected. Most miss the target—on its trip across Japan, the neutrino beam expands to a cross-sectional area almost a million times the size of Super-Kamiokande. Only a few per trillion of the neutrinos that do pass through the detector undergo a weak interaction there, and only a few percent of those are electron neutrinos, the valuable products of the flavor oscillation T2K seeks to measure. All told, it takes some 10^{21} accelerator protons to produce the few dozen electron neutrinos that the experiment has detected.

Super-Kamiokande can distinguish muon neutrinos from electron neutrinos, but not neutrinos from antineutrinos. Fortunately, that ambiguity can be resolved at J-PARC's end. The pions produced by the accelerator protons come in both positively and negatively charged varieties, with π^+ decaying into antimuons and muon neutrinos, and π^- decaying into muons and muon antineutrinos. By steering either π^+ or π^- into the decay volume, the researchers can produce a nearly pure beam of either neutrinos or antineutrinos. Since 2014, T2K has been alternating between neutrino and antineutrino modes, with approximately the same number of accelerator protons spent on each.

Symmetry violation

Even without any *CP* violation, T2K's data for neutrinos and antineutrinos aren't expected to be identical. The neutrinos propagate through rock made of matter, not antimatter, and they're measured by a detector made of matter. Both of those so-called matter effects bias the relative probabilities of neutrinos and antineutrinos oscillating and being detected. Furthermore, the oscillation and detection rates depend on the neutrino's kinetic energy, which can't be directly controlled or measured; it can only be inferred from the energy of the Cherenkov radiation in the detector through a complicated nuclear-physics calculation.

Once those effects are taken into account, T2K's observations—90 electron neutrinos and 15 electron antineutrinos over 10 years—suggest not only that muon neutrinos transform more readily than muon antineutrinos do but that the imbalance is close to as large as it could possibly be. Theory predicts that if δ_{CP} were $-\pi/2$, the value that produces maximal *CP* violation in favor of antineutrino oscillation, the experiment should have seen 82 electron neutrinos and 17 electron antineutrinos. A *CP*-conserving δ_{CP} of 0 would yield 68 electron neutrinos and 20 electron antineutrinos. And a δ_{CP} of $\pi/2$ would result in 56 electron neutrinos and 22 electron antineutrinos.

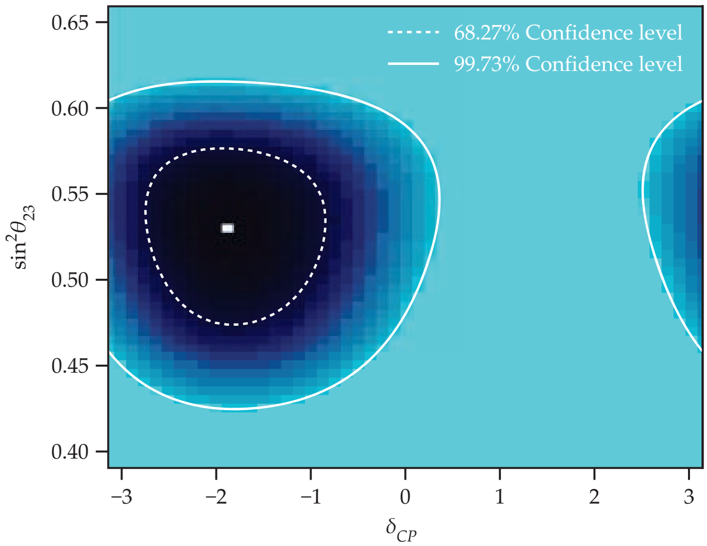


FIGURE 2. CONFIDENCE LIMITS on neutrino oscillation parameters derived from the T2K experiment's data. The plot shows constraints on the mixing angle θ_{23} and the *CP*-violating phase δ_{CP} at the 68.27% (1σ) and 99.73% (3σ) confidence levels; the result marks the first time that any possible values of δ_{CP} have been excluded with 3σ confidence. (Adapted from ref. 1.)

The small numbers and complicated physics make for large uncertainties. Figure 2 shows the confidence limits at 1σ (68.27%) and 3σ (99.73%) on δ_{CP} in conjunction with θ_{23}, another parameter that affects the oscillation probability. The data are most consistent with a negative δ_{CP}, and most of the positive values are excluded with 3σ confidence.

That analysis assumes the so-called normal mass order, in which m_3, the neutrino mass that differs the most from the other two, is the largest of the three masses. The inverted order, in which m_3 is the smallest, predicts different oscillation dynamics that are less consistent with the data. The true order of the neutrino masses remains an open question, but the T2K results make the normal order look a bit more likely.

Next generation

T2K isn't the only accelerator experiment looking for signs of neutrino *CP* violation. There's also NOνA (NuMI Off-Axis ν_e Appearance), which generates muon neutrinos at Fermilab in Illinois and detects electron neutrinos in northern Minnesota. The NOνA collaboration's most recent results, published last year, also favor the normal over the inverted mass order.[2] Although the NOνA data are best fitted by a δ_{CP} of 0, they can't yet exclude any δ_{CP} values with even 1σ confidence, and they're compatible with T2K's findings.

Both the T2K and NOνA experiments are ongoing, and they'll continue to collect data in the coming years to pin down δ_{CP} and other neutrino properties. At the same time, the next generation of experiments is in the works. In Japan, the planned Hyper-Kamiokande detector, a water tank even larger than Super-Kamiokande, will allow data collection 20 times as fast as T2K currently achieves. And in the US, DUNE (Deep Underground Neutrino Experiment) will pair neutrinos generated at Fermilab with a liquid-scintillator detector in western South Dakota.

Both the new experiments are due to come on line in 2026 or 2027, although that schedule reflects the plan before the outbreak of COVID-19. Once they're up and running, how long they'll take to get definitive results depends on what δ_{CP} turns out to be. A large symmetry violation is easier to spot than a small one. If *CP* violation is maximal, as the T2K results suggest it might be, DUNE and Hyper-Kamiokande could reach a 5σ measurement in as little as two years of operation.

Johanna Miller

References

1. T2K collaboration, *Nature* **580**, 339 (2020).
2. M. A. Acero et al., NOνA collaboration, *Phys. Rev. Lett.* **123**, 151803 (2019).

Designer proteins act as logic gates

Controllable protein–protein interactions modify the behavior of a yeast cell and a human T cell.

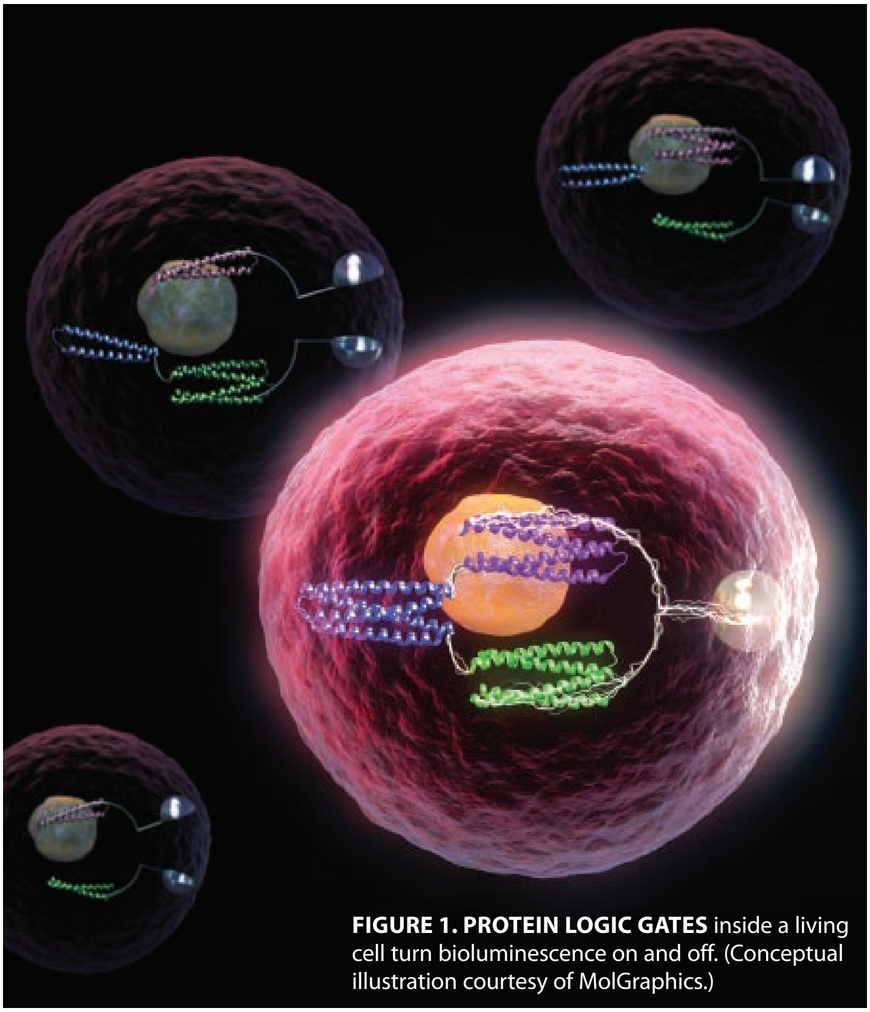


FIGURE 1. PROTEIN LOGIC GATES inside a living cell turn bioluminescence on and off. (Conceptual illustration courtesy of MolGraphics.)

What if the behavior of natural and synthetic cells could be programmed like computers? Then a programmer could turn cellular behaviors on and off in a living organism by adding certain molecules. Cells already sense and respond to stimuli. But with external control, those behaviors could be put to work in medical or biotechnological tasks—for example, as smart drug delivery or telling bacteria to clean up toxic waste. To that end, researchers have engineered logic gates, largely using DNA and RNA, that introduce programmable circuitry into cells.

Although those nucleic-acid logic gates are easy to program, cells make decisions through protein–protein interactions. A protein-based logic gate can speak directly to a cell's existing decision-making circuits. Researchers have already modified naturally occurring protein signals to introduce new logic pathways. But those implementations are inherently limited in scope by the number, properties, and geometries of the proteins.

Now David Baker of the University of Washington in Seattle and his colleagues have designed and constructed proteins from scratch that can be modularly assembled to perform an array of logic operations both inside cells,[1] as represented in figure 1, and outside of them.

De novo design

The study's lead author, Zibo Chen, was interested as an undergraduate in engineering protein–protein interactions, which underlie much of natural cellular decision making. As synthetic biology becomes more prevalent and complex, protein–protein interactions will need to be designed deliberately. So when he joined Baker's lab in graduate school, Chen pursued tunable protein interactions using *de novo* proteins—that is, proteins designed and built from the ground up.

A protein's particular sequence of amino acids reliably folds into the shape that minimizes energy and balances the attraction and repulsion among amino acids and the fluid surrounding them. The difficulties for researchers who design proteins from scratch are calculating the energy accurately and sorting through the large set of possible sequences: In an average-length protein, the 20 amino acids can form 20^{200} possible sequences, of which only 10^{12} occur naturally.

In the 1990s Baker and his group developed a program called Rosetta that uses a Monte Carlo sampling algorithm to solve for the lowest-energy folded structure of a given protein's amino-acid sequence. They eventually started tackling the inverse problem—selecting a three-dimensional structure and then finding the sequence that produces it. The problem is complicated; it often takes a combination of known peptide fragments and an iteration between predicting a sequence and double-checking what structure it yields. To meet the huge computational load, Baker founded a project called Rosetta@home, in which people can offer their personal computers for protein computations. (At the moment the project is modeling SARS-CoV-2 proteins and small proteins for potential therapeutic and diagnostic uses.)

In current *de novo* protein design, researchers often start by picking a desired function or shape and then finding a suitable polypeptide backbone structure. The structure needs to have a high chance of being the lowest-energy state, or it won't form in practice, and it needs to accommodate a core of amino acids. One common backbone is a bundle of helices, whose optimization is manifest in equations developed in 1953 by Francis Crick for describing how helices pack together. Once researchers select a backbone shape and solve the inverse problem for the corresponding amino-acid

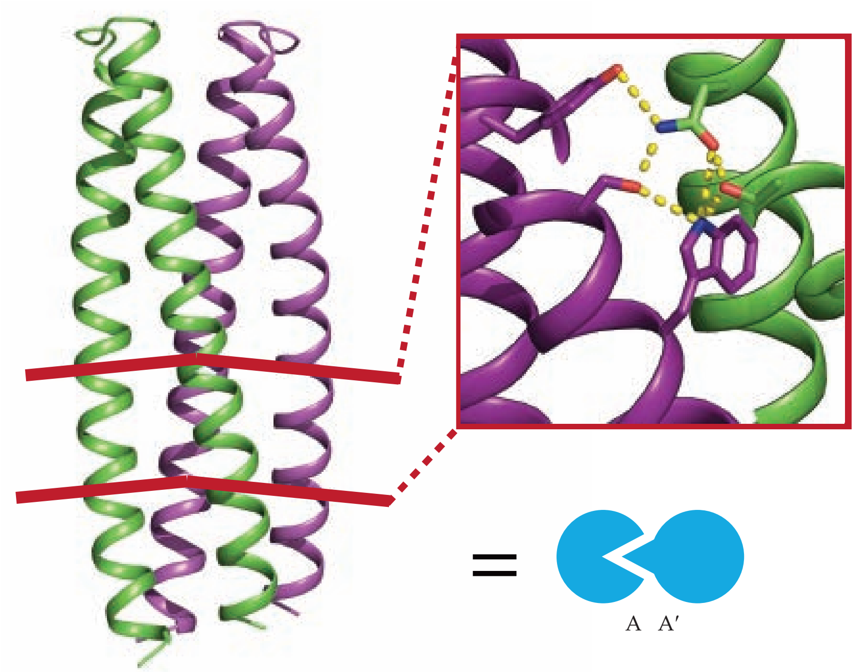


FIGURE 2. TWO DESIGNER PROTEIN MONOMERS A and A′ (green and purple coils) form a dimer through a network of complementary hydrogen bonds, as shown in the inset. The heterodimer they form, A:A′, is represented by the interlocked symbols shown in blue. (Adapted from ref. 1.)

sequence, they perform a similar sampling process and inverse problem for the core amino acids. Finally, they manufacture or buy synthetic DNA that encodes that sequence. With that synthetic DNA, *Escherichia coli* bacteria produce the designed proteins.

The algorithms that design *de novo* proteins optimize for stability, and the resulting proteins are stable—often too stable. The protein–protein interactions necessary for logic-gate applications generally happen at interfaces that are energetically perturbed, and stable proteins resist perturbation. Protein design thus needs to strike a balance between stability and functionality.

Logical proteins

Logic gates sense and respond to inputs in a set way. To create a suite of biological logic gates, their building blocks should ideally have similar, modular structures and come in mutually orthogonal pairs. DNA is a prime example; only specific

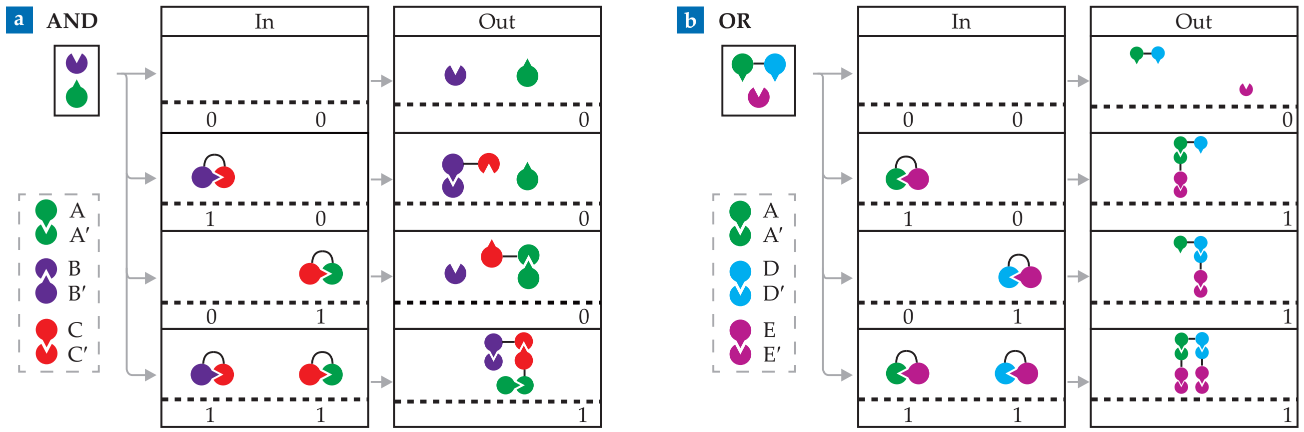


FIGURE 3. AND AND OR GATES can be constructed from combinations of isolated and linked protein monomers. **(a)** An AND gate can be made from three pairs of complementary monomers—A:A′, B:B′, and C:C′. In the presence of A and B monomers, complexes of A′ and C monomers (A′–C) and C′ and B′ monomers (C′–B′) serve as inputs. A and B are joined, an output signal of 1, only if all monomers form a single complex, as in the bottom row. **(b)** Surrounded by E′ monomers and A–D complexes, the inputs A′–E and D′–E produce a single output complex —a signal of 1—when at least one input is present, as in an OR gate. (Adapted from ref. 1.)

nucleotides pair together, and they all fit in a double-helical backbone. Using DNA as inspiration, Baker and his colleagues designed a collection of proteins that all took a similar form: a bundle of coiled backbones with a network of hydrogen bonds, as shown in figure 2, that arose from polar amino acids and provided a bonding interface. Two monomers fitted together like a lock and key, such that they formed a dimer only if they had complementary interfaces that create a network of hydrogen bonds.

In 2016 Scott Boyken, then Baker's postdoc, introduced a computational method to enumerate all the possible hydrogen bond networks for a given backbone structure. Boyken, Chen, and their colleagues then employed the technique to design a set of 39 orthogonal protein pairs.[2] With more pairs than, say, DNA's two pairs of nucleotides, gate complexity has fewer limits. Because the networks were designed with atomic-level accuracy, far more than those 39 pairs are possible.

Turning protein pairings into logic gates takes judicious combinations of complementary monomers. For example, take three pairs of monomers that form the heterodimers A:A′, B:B′, and C:C′. With complexes of A′ and C monomers (A′–C) and C′ and B′ monomers (C′–B′) as inputs, the system acts as an AND gate, as shown in figure 3a. If both inputs are present in a sea of A and B monomers (bottom row), all the monomers end up in a single chain, and an output signal of 1 results. Otherwise, the A and B monomers end up in different molecules, and there is no signal.

The signal from those fully bonded complexes depends on what researchers bind to the noninput monomers. For example, Baker and his colleagues constructed an AND gate in a yeast cell containing a transcription protein with two separable parts. They fused one part to the A monomer and the other part to the B monomer. In the positive output configuration, those two parts are close together, and that positioning activates a gene responsible for increasing the yeast's growth rate. Otherwise, the growth rate stays the same.

In similar fashion to the AND gate, an OR gate can be devised, as shown in figure 3b, from three pairs of monomers, A:A′, D:D′, and E:E′. The inputs are fused A′–E and D′–E in a sea of E′ and A–D. A signal occurs any time one or both inputs combine the other monomers into a single complex—that is, so long as there is at least one input. The researchers took similar strategies to create NAND, NOR, XNOR, and NOT gates with two and three inputs.

Cellular output

To work in a cell, the logic gate can't be too sensitive to population imbalances in the available monomers. So monomers must be tied in a suitable way, such as with an optimized length of linker and with the binding interfaces hidden. For two properly fused monomers, the energy required to expose their bonding sites is provided by the sum of the binding energies for their complementary monomers—for example, the A′–E complex unfolds only if both A and E′ are present; if only A is present, the energy barrier is too high, and A′–E won't unfold. Partial complexes won't form, and a major imbalance in the inputs won't throw off the gate's function.

To test that cooperativity in actual samples, the researchers used native mass spectrometry, which measures the populations of different compounds in their nearly native state. They found that even a sixfold imbalance in the inputs didn't affect the number of fully bonded complexes.

Biological logic gates could be important for medical treatments. As a demonstration, the researchers used human T cells, essential to immune responses. When T cells fight chronic infections and cancer, they often suffer from a dysfunction known as exhaustion, in which there's a sustained inhibitory signal that stops T cells from doing their job.[3] But T cells need transient inhibitory signals to prevent autoimmune disorders. By modulating their logic-gate inputs, Baker and his group were able to selectively repress a gene thought to modulate exhaustion. With their logic gate, T cells may be able to overcome inhibitory signals only in the case of exhaustion.

Heather M. Hill

References

1. Z. Chen et al., *Science* **368**, 78 (2020).
2. Z. Chen et al., *Nature* **565**, 106 (2019).
3. E. J. Wherry, M. Kurachi, *Nat. Rev. Immunol.* **15**, 486 (2015).

Transportable atomic clocks achieve laboratory precision

When deployed in field-based experiments, the devices could improve timekeeping standards and test fundamental physics.

The 1967 General Conference on Weights and Measures defined the SI unit of time, the second, based on an atomic transition—specifically, between two hyperfine levels of the ground state of cesium-133. (See PHYSICS TODAY, August 1968, page 60.) Although Cs atomic clocks remain the standard, their time might be running out. Their underlying atomic transition is excited by radiation with a microwave frequency around 9×10^9 Hz, and after decades of advances, a Cs clock's frequency can be measured with a fractional uncertainty $\Delta\nu/\nu_0$ of about one part in 10^{16}.

But clocks based on optical transitions operate at frequencies around 10^{14} Hz, which gives them an advantage in the push for lower uncertainty. (See the article by James Bergquist, Steven Jefferts, and David Wineland, PHYSICS TODAY, March 2001, page 37.) The current record, 9.4×10^{-19}, was set in 2019 by an aluminum ion–based optical atomic clock at NIST.

In some applications, optical clocks can't yet provide any practical benefit over their microwave counterparts.[1] For example, atomic clocks around the world are regularly compared with each other to maintain International Atomic Time. Those comparisons are done via satellite intermediaries, but the clocks on those satellites use microwave transitions, so even if the Earth-based laboratory clocks were more precise, their comparisons wouldn't be. The comparisons also rely on precise geodetic measurements that can be difficult and time-consuming to obtain. According to general relativity (GR), gravity slows the passage of time, so a clock at sea level ticks more slowly than one on a mountaintop. Atomic clocks don't need a mountain to register that difference: For clocks with a precision level of 10^{-18}, even a few centimeters matters.

Small, transportable optical clocks could replace Cs ones for high-precision applications. Mounting them on satellites would facilitate worldwide clock synchronization and improve GPS accuracy, and networks of optical clocks could measure geopotential differences with centimeter-level precision. They would be a valuable tool to test Lorentz invariance and search for dark matter. But the frequency uncertainty in transportable optical clocks has lagged behind that of lab-based devices.[2] The tradeoff is between portability and precision: The best timekeepers are laboratory-based atomic clocks that rely on large, heavy equipment like optical tables to create well-controlled, mechanically isolated environments.

The pair of optical clocks shown in figure 1 have now achieved a fractional uncertainty of 5×10^{-18} while operating outside the lab—an order of magnitude better than previous transportable clocks.[3] The devices were developed by Masao Takamoto and Noriaki Ohmae at Japan's RIKEN research facility, Ichiro Ushijima and Hidetoshi Katori at the University of Tokyo, and their colleagues at other Japanese institutions.

Keeping cool

Optical atomic clocks fall into two categories: optical lattice and single ion. Each has pros and cons. Lattice clocks rely on measurements of many atoms for precision, and they have the potential to outperform single-ion clocks. However, atoms in lattice traps are more sensitive to electric field perturbations—from the trapping lasers, charges on nearby surfaces, and ambient blackbody radiation (BBR)—than those in ion traps. The energy-level shifts caused by those perturbations can obliterate performance gains over not just single-ion clocks but also clocks with microwave-range transitions.

Takamoto, Katori, and colleagues demonstrated the first optical lattice clock at the University of Tokyo in 2003. Strontium was a convenient choice because the energy levels for its clock transition and for laser cooling are excited by diode lasers. Since then, the researchers have refined their optical lattice clock. For example, they improved the stability of clock comparisons by rejecting the noise from the clock laser, and they precisely determined the conditions under which the lattice lasers would least disturb the Sr atoms' energy levels. Those advances and others were incorporated into the transportable clocks.

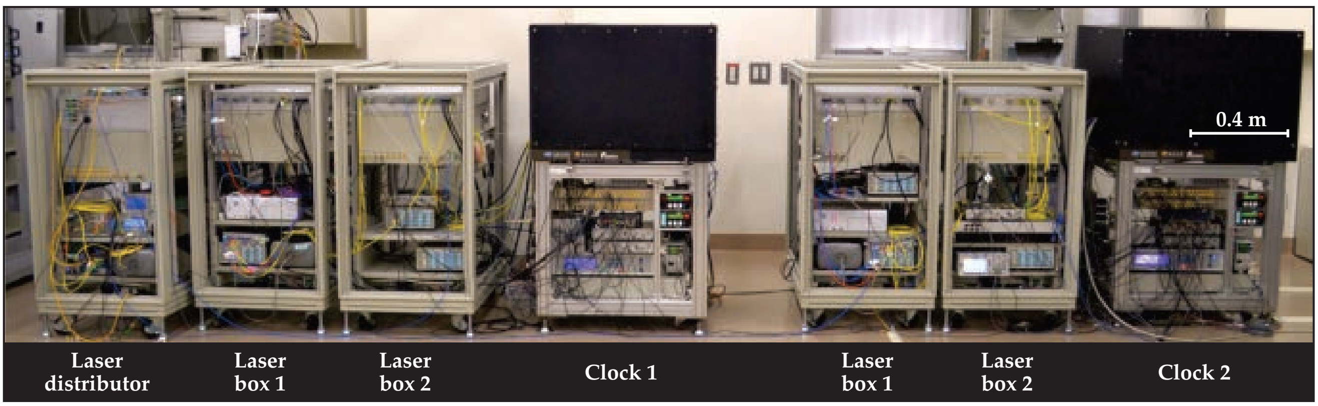


FIGURE 1. TWO TRANSPORTABLE OPTICAL ATOMIC CLOCKS carry spectroscopy chambers (black boxes) and supporting equipment. Each clock also has two laser boxes. Box 1 includes three lasers, two for cooling and one for repumping. Box 2 encloses lasers for creating an optical lattice trap, further cooling, and exciting the atoms for timekeeping. The two clocks are connected by a telecommunications fiber through the laser distributor. (Adapted from ref. 3.)

To protect the Sr atoms from stray photons, the RIKEN researchers developed BBR shields, which they have installed in their clocks since 2015. Then as now, the first step in each measurement of the clock transition is to cool Sr atoms down to a few microkelvin and load them into a one-dimensional optical lattice trap formed in a ring cavity (figure 2). A pulse from one of the two lattice lasers then nudges the trapped atoms into an 18-mm-long temperature-controlled chamber—the BBR shield—whose inner walls are painted with a high-absorbance black coating to prevent any stray photons from bouncing around.

Inside the BBR shield, the Sr atoms undergo a final cooling step before having their clock transitions measured. The device keeps time by probing the frequency of radiation corresponding to the $^1S_0$–$^3P_0$ transition, but it doesn't measure that frequency directly. Rather, the frequency of the clock laser is tuned to match the Sr clock transition as closely as possible. The laser is directed at the atoms, and the more closely it matches the clock transition, the more atoms it excites. The trap then carries the atoms back out of the chamber so the fraction of excited atoms can be measured, and the clock laser's frequency is adjusted to find the maximum.

Improving precision in a lab-based clock is one thing; maintaining that precision in a transportable clock is another. "Four or five years ago, we were just happy with clock comparisons at 10^{-18} using laboratory-based machines," says Katori. "At that point it was possible to think about the experiment, but it was technically too hard to do." Making the clocks compact and stable enough to leave the lab required specialized, maintenance-free equipment. The lasers had no adjustment knobs, and the optical components were welded in place. The researchers also collaborated with the Shimadzu Corp to develop electronic devices, such as laser controllers and oscilloscopes, without the bulky control panels and displays of commercial equipment. All the clocks' components were controlled remotely through a single personal computer.

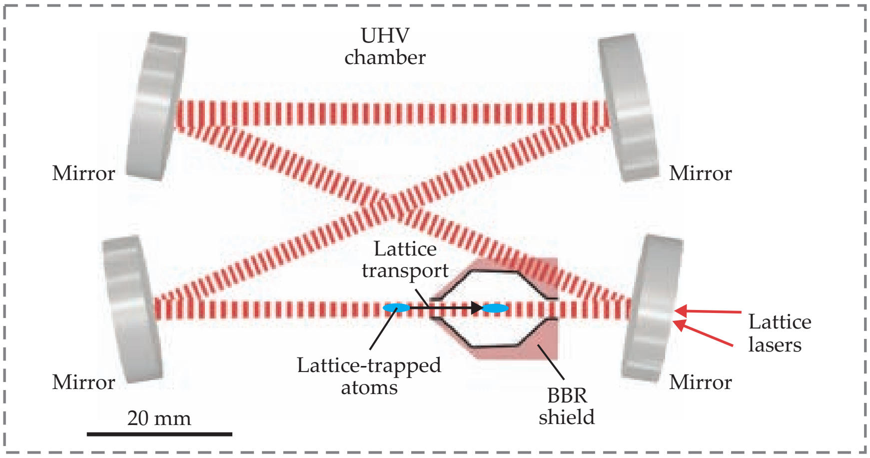


FIGURE 2. THE SPECTROSCOPY CHAMBER for each transportable strontium atomic clock contains a ring cavity that creates a one-dimensional optical lattice trap. Once a collection of atoms is trapped, a pulse from one of the lattice lasers transports it into a blackbody radiation (BBR) shield that protects the atoms from stray IR photons and strontium atoms. While inside the chamber, the atoms undergo an additional cooling step before excitation. (Adapted from ref. 3.)

Onward and upward

One use of atomic clocks is to precisely measure a gravitational redshift—the frequency difference between two identical clocks at different gravitational potentials. Detecting the difference isn't particularly difficult; GPS satellites adjust their times by 38 μs every day to account for relativistic effects (see the article by Neil Ashby, PHYSICS TODAY, May 2002, page 41). But precise measurements of gravitational redshifts can rigorously test GR's predictions. Some more complete descriptions of the universe require modifications to GR to account for, say, dark energy or the unification of gravity with the other fundamental forces. Measuring a deviation—or lack thereof—from GR's predictions would help point theoreticians in the right direction.[4]

The fractional frequency shift between two clocks is related to their gravitational potential difference ΔU by $\Delta\nu/\nu_1 = (1 + \alpha)\,\Delta U/c^2$ where ν_1 is one of the clock frequencies and c is the speed of light. If GR is correct, α is exactly zero. Tests of GR's gravitational redshift predictions try to establish the value of α as accurately as possible, and they're facilitated by two parameters: a large gravitational potential difference and a precise measurement of the resulting frequency difference.

As a demonstration of their clocks, Takamoto, Ohmae, and Ushijima took the clocks to the Tokyo Skytree broadcasting tower to measure a gravitational redshift. They placed one clock at the tower's base and brought another up to the 450-m-high observatory floor. The height difference between the clocks was established with centimeter precision by using navigation satellites and laser ranging, and a pair of gravimeters determined the clocks' local gravitational accelerations. Putting that together with the frequency measurements from the two clocks, the researchers calculated a value of $\alpha = (1.4 \pm 9.1) \times 10^{-5}$. It's the best constraint on α from a ground-based measurement and is nearing the limit established by space-based experiments done using satellites separated by thousands of kilometers.[5]

The Skytree tower proved to be a challenging environment for the clocks because vibrations from nearby trains were unexpectedly large. The clock laser is particularly sensitive to noise, and even after the researchers added active vibration isolation, the vibrations limited the precision of the frequency comparison between the clocks. Although the researchers could have chosen a more amenable environment, they thought it was important to develop an optical clock with the ability to perform in adverse conditions. Katori sees it as a surmountable challenge: "By developing and installing a more stable laser system in the future, we will be able to significantly improve the stability of the clocks."

Christine Middleton

References

1. F. Riehle, *C. R. Phys.* **16**, 506 (2015).
2. J. Cao et al., *Appl. Phys. B: Lasers Opt.* **123**, 112 (2017); S. Origlia et al., *Phys. Rev. A* **98**, 053443 (2018).
3. M. Takamoto et al., *Nat. Photonics* (2020), doi:10.1038/s41566-020-0619-8.
4. See, for example, A. Derevianko, M. Pospelov, *Nat. Phys.* **10**, 933 (2014).
5. P. Delva et al., *Phys. Rev. Lett.* **121**, 231101 (2018).

PT

Universities overcome bumps in transition to online teaching

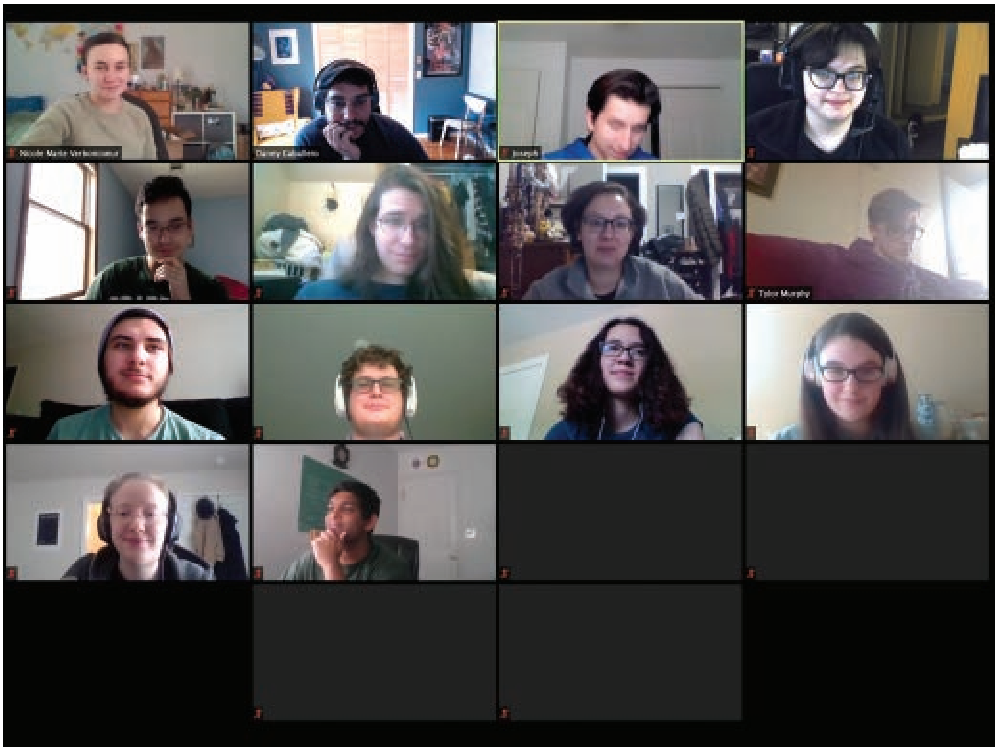
Instructors grapple with how to administer exams that meaningfully assess students, suppress cheating, minimize anxiety, and preserve privacy.

Teachers at universities worldwide are catching their breath as the first term in mass online teaching wraps up. The shuttering of campuses when social distancing was implemented to slow the spread of COVID-19 set off a scramble to deliver college education remotely. Faculty had to move their courses online, work from home, and engage students who had varying external distractions and uneven internet access. The difficulties of the transition—including the thorny issue of exams—cut across all subjects, but huge introductory classes and laboratory instruction pose particular challenges in physics.

"Our goal was to provide all components of instruction, even while satisfying constraints and accommodating students who have challenges at home," says Brian DeMarco, associate head for undergraduate programs in physics at the University of Illinois at Urbana-Champaign. Courses that serve engineers still have to meet the certification requirements for engineering, he says, and in the US all courses must comply with FERPA, the federal Family Educational Rights and Privacy Act. For example, graded assignments have to be returned to students via secure systems, not by email; and students cannot be identified in publicly accessible videos.

"Students and faculty had the rug pulled out from under them," says Vernita Gordon, who took on a coordinating role for remote teaching in the University of Texas at Austin physics department. Like many professors, she is also caring for her children, whose school is closed due to the pandemic. And faculty also have to deal with overtaxed Wi-Fi bandwidth: Another instructor in Austin recounted how he had lectured live to a dead connection; when he got back online he had to repeat 10 minutes of his lecture. "Everyone is doing their best," says Gordon. "But it's much harder for students to learn, and it's much harder to know what they have learned."

DANNY CABALLERO



STUDENTS IN DANNY CABALLERO'S UPPER-DIVISION electricity and magnetism class tune in via Zoom.

Many physics departments were already at varying stages of putting parts of courses online before the COVID-19 outbreak. Chris Waltham chairs undergraduate studies in the University of British Columbia's physics department. "We had all of the tools in place," he says. "I've been amazed at how seamless the transition has been." Still, he says, one concern is low attendance, "although most of the class appears out of the woodwork for quizzes."

A mad scramble

Faculty members were generally given great leeway for how they transitioned to online teaching. Decisions about live versus recorded lectures, video platforms, how to run labs, and how to administer exams were largely left to individuals or team-teaching cohorts. Instructors submitted their teaching plans to university administrators. Many universities provided training in Zoom or other online platforms, and some physics departments have employed in-house technical help. Some institutions ponied up for electronic writing pads, document cameras, and other tools that teachers need for makeshift distance instruction.

"We had to figure out how to use webcams and microphones—anything we could get our hands on to start recording lectures at home," says Michael Dubson, associate chair of physics for undergraduate studies at the University of Colorado Boulder. "It was bumpy, but my colleagues and I stayed in touch and we all got things to work."

Approaches to distance teaching vary

by instructor style, class size and level, and student needs. "The way people have presented material in class may or may not translate well to the online context," says Gordon. Before the COVID-19 outbreak, her honors physics class for nonmajors was very interactive: "We would break into small groups, and a TA and I would circulate and talk to them," she says. "We can replicate that format to some degree on Zoom, but it's not the same. I feel the loss of the personal interactions pretty strongly."

Many instructors complain that breakout rooms on video-conferencing platforms hamper interactions both among students and between students and instructors. Physicist Carl Wieman of Stanford University says that getting comfortable with online teaching requires a "learning curve," but that if breakout rooms are used in groups of just a few students who have clear instructions and goals "there can be a lot of interaction." In fact, he says, for large classes the interaction could be better than in-person.

At Georgia Tech, the guiding principle has been to stick as closely as possible to the original course, says Edwin Greco, the lead instructor for a calculus-based introductory physics course with about 900 students. He and his colleagues have chosen to continue delivering their lectures at the scheduled time. They later upload videos of their lectures to allow students who are home in faraway time zones, encounter failures in internet connectivity, or have other conflicts to keep up with class on their own schedules.

During in-person lectures, says Greco, the instructor would pose a question every 10 minutes or so. The students would discuss the question with their neighbors for a few minutes and then submit their answers. "If most of them get the right answer, I move on," says Greco. "If not, I adjust the live lecture." That doesn't work as well online: Student discussion is harder to facilitate and web-based interactions are much slower. In person, he adds, "you can tell if someone is paying attention, but that's hard to do virtually."

Other instructors chose to teach asynchronously, sometimes in a flipped mode, with students watching lectures before attending virtual discussions. Some instructors, including Dubson, embed questions in their video lectures such that students can't continue until they commit to an answer. "This allows us to require that they think," he says.

Danny Caballero at Michigan State University is teaching a senior-level electricity and magnetism (E&M) class with 24 students. The class is small enough for him to stay in touch with his students, and his main aim is that they demonstrate understanding of the material. After in-person collaborations ceased, he had the students write and solve quizzes and review each other's work. Logistically, he says, the transition has been easy for him.

For many students, though, the transition has been tough, Caballero says. "They are taking three, four, or five classes online. They have varying financial situations—they have lost their campus jobs, it's not clear they all have food security." Anxiety among students is a big issue, he says. "Some are isolated, some are depressed."


KAREN ZUKOR

JOEL FAJANS SOLDERS COMPONENTS for kits with microcontrollers, resistors, and other circuit parts for an advanced laboratory class at the University of California, Berkeley. Fajans delivered the kits to his 55 students, who had scattered over three continents when COVID-19 shuttered the university. From home, the students built amplifiers and measured and analyzed noise to obtain Boltzmann's constant.

Hands-off experiments

For most North American campuses, the term was at least half over when the lockdowns began. In lab courses, students had generally performed half or more of the experiments. Douglas Bonn had the roughly 100 students in his second-year lab course at UBC shift their emphasis to communications skills and writing.

In other classes, students switched to simulations, such as the free, interactive PhET experiments developed at CU Boulder in the early 2000s. The experiments in physics explore pendulums, Snell's law, gas density, circuit construction, and more; the full library includes simulation experiments in math, biology, chemistry, and Earth science.

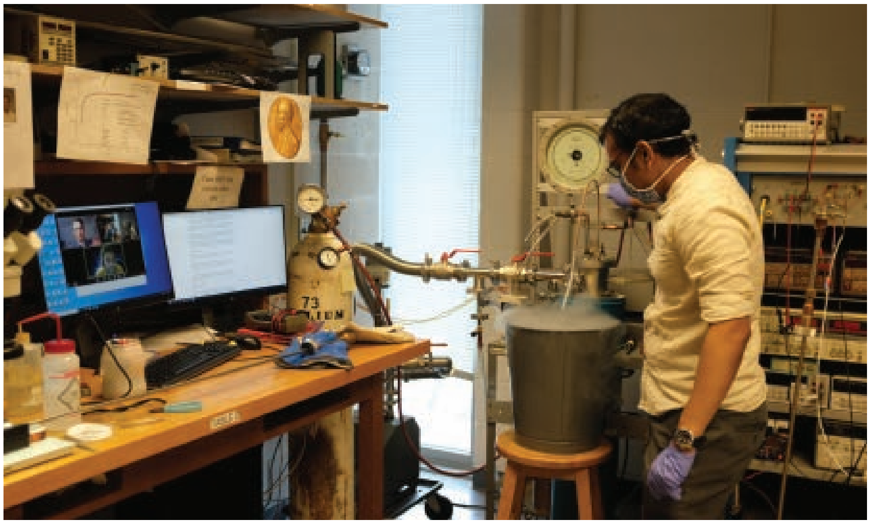
At Illinois, teaching assistants went to the physics building to perform experiments in real time with undergraduates, who could partially run the measurements via Zoom. Another option is for students to do experiments from home with a smart phone or an iOLab, a smart phone–sized device that faculty at Illinois developed a few years ago, with built-in sensors that measure force, pressure, temperature, and other quantities. The pandemic has made the iOLab so popular that the manufacturer, Macmillan Learning, may not be able to keep up with demand, says Dubson. For some courses, instructors have created kits for students to perform experiments at home.

Many upper-division lab classes, however, require that students "get their hands on the equipment," says Bonn. "That's an interesting challenge, and we may delay those courses."

Assessment, cheating, and stress

Exams, especially at the introductory level, are perhaps the trickiest and most controversial aspect of the move to remote teaching. Disagreements over how to handle them have strained relations in some departments. "How can we be equitable to students? How do you avoid biasing against those who don't have good internet access?" says a state-university physics professor who did not want to be identified. "It boils down to the balance between suppressing cheating and meaningfully evaluating students." The

BRIAN DEMARCO



PHYSICS GRADUATE STUDENT SHUBHANG GOSWAMI measures superconductivity in thin films. The experiment is part of a senior-level lab course at the University of Illinois at Urbana-Champaign. Undergraduates participate remotely and can partially control the measurements.

decisions are “heartbreaking,” says the professor. “I have grave doubts that we are on the right path.”

In the large freshman E&M class Dubson team-teaches at CU Boulder, the third exam of the semester was administered online, with no security measures. “We have bits of evidence that significant cheating was going on,” he says. For the final exam, Dubson and his colleagues considered several options. One was proctoring software. But between students saying they lacked webcams and known problems with the software, the instructors nixed the idea. They considered giving students varying selections from a large bank of questions. “That approach is probably the most fair and least stressful for students,” says Dubson. “But it’s very time consuming for the faculty to put together.”

In the end, Dubson’s team went with a format in which students submit the answer to one problem before they can see the next one, and they can’t revisit problems. The order of questions was randomized, says Dubson. “That makes it nearly impossible for students to collaborate during an exam, but it does not prevent a student from hiring an impostor to take an exam in their place.” The approach is unpopular with both students and faculty, he says. “No one regards this as a reasonable test-taking environment. Students can’t ponder questions and budget their time. It increases the stress on students. But most students accept the need for some exam security measures.”

The University of Illinois at Urbana-Champaign has contracted with an online proctoring company, but the physics faculty have decided not to use it, says Timothy Stelzer, a high-energy theorist who is team-teaching introductory calculus-based E&M for engineers. The class has 600 students. “We didn’t want to invade privacy,” he says. “I’m not convinced that proctoring solves the problem [of cheating], and it adds a lot of stress.” The instructors extended the time for the final exam from 90 to 120 minutes to accommodate slow internet connections, and they offered the exam at different times. In addition, the first question requires students to agree to an honor code; only then can they see the actual test.

UT Austin’s Gordon also included an oath of honor. And she opted for proctoring software that locks browsers, records clicks, and uses artificial intelligence to monitor student movements. “It’s the best in a set of unsatisfactory solutions,” she says. After using the software for a midterm exam, it seemed “much less invasive than I feared,” she says. She watches the students via webcam only if the software identifies suspicious behavior.

The exam issue is thorniest for lower-division courses. At higher levels, classes are smaller, and faculty are more comfortable assessing students with open-book exams or projects.

But Wieman says open-book exams can work at all levels. They allow for more meaningful questions and are better tests of realistic problem-solving capability, he says. At Stanford, where many of his colleagues are choosing the open-book route, exams turn into a learning experience "with far less artificial hoop-jumping and guessing at instructors' idiosyncrasies," he says.

Many universities have adjusted their policies on course withdrawal and grades. "A lot of students come in well-prepared," Stelzer says. "But there is a population for whom that is not true, and that tends to be the same people who don't have access to the internet and have a harder situation at home—sharing computers, occupying cramped spaces, or taking on extra family responsibilities." To accommodate such inequities, and to be sympathetic, he says, many instructors have loosened deadlines for labs and quizzes. And most universities have extended the deadlines for dropping classes to just before—or even after—final exams.

Many universities have ditched grades this term in favor of pass/fail. A few, such as Georgia Tech, have retained grades, despite student complaints. And many institutions are giving students the choice of either a grade or a pass/fail. At UBC, the science dean issued a rare decree requiring faculty to calculate grades with two different weightings for the final exam—30% and 5%; students will receive the better grade. Or they can opt for pass/fail. In late March, the American Physical Society sent a letter to department heads urging their graduate admissions committees to treat this term's grades "holistically."

Overall, the wholesale transition to remote teaching created a mad scramble and a lot of improvisation. But many faculty say they've learned things they'll take forward for future online teaching and for when in-person classes resume. Jonathan Wurtele of the University of California, Berkeley, notes that his campus occasionally closes due to smoke from nearby fires. "We will put the knowledge of remote teaching to use in the future," he says. Similarly, Karen Daniels of North Carolina State University says she'd be comfortable teaching remotely for a day or so if she leaves town to attend a conference. But, she says, "even if we have found replacements for all the parts of a normal face-to-face class, it's not the same. We are not delivering what we need to."

Online office hours, for which students choose a time slot for a video conference, could continue to work well especially for commuter students, according to several instructors. Many professors found that students were good at helping each other in the text chat boxes in videoconferencing software, and they hope to incorporate that type of help in their in-person courses. Andrew Loveridge of UT Austin notes that with the transition to remote teaching, "we are forced to think about every part of our courses. Nothing will survive on its own inertia."

Toni Feder

COVID-19 pandemic modeling is fraught with uncertainties

Policymakers face a plethora of predictions on how the disease will proceed and when it might resurge.

A self-described optimist, Pinar Keskinocak doesn't like to be the bearer of bad news. But the model she co-developed at Georgia Tech of the COVID-19 pandemic in that state paints a "really bleak" picture of what lies ahead when physical distancing slowly erodes after shelter-in-place and stay-at-home orders end.

The model, which forecasts the outbreak in Georgia at the census tract level—county subdivisions that average 4000 inhabitants—shows that even if lockdowns had been extended through mid-May instead of being lifted 1 May, the rate of new infections would come roaring back once people returned to their daily routines.

Georgia was one of the first states to end shelter-in-place orders and permit some businesses to reopen. Although continued adherence to social distancing guidelines will tamp down the state's peak numbers of new infections, even strict compliance—including the voluntary quarantining of all persons in households where only one member is infected—won't prevent the outbreak from surging to levels far higher than those yet experienced. The real peak of new cases in Georgia, says the model, is predicted to come in June or July (see graphs on pages 26 and 27) and potentially overwhelm health-care facilities in some parts of the state.

The Georgia Tech model's findings, which were shared with state government officials—Keskinocak won't say exactly whom—before governor Brian Kemp's decision to lift stay-at-home restrictions, presented a different portrait of the pandemic from the widely reported modeling results coming from the University of Washington's Institute for Health Metrics and Evaluation. Until 26 April, after which it was significantly revamped, the IHME model had predicted that the peak daily death toll from COVID-19 in Georgia had already passed, even before the mid-April zenith it had forecast for the nationwide death rate. On 3 May, a new, hybrid version of the IHME model was projecting that daily deaths in Georgia would peak on 30 May, well after the forecasted 1 May peak in daily US deaths.

Another model, developed at Los Alamos National Laboratory (LANL), reported with 96% confidence as of 3 May that the daily rate of new cases in Georgia has peaked. Although LANL's model doesn't explicitly include the effects of interventions such as sheltering in place and social distancing, it assumes that some social distancing measures will continue through the forecast period. LANL modeler Dave Osthus says the model won't be adjusted to account for the ending of lockdowns because the extent to which people will actually change their behavior is unknown.

Many other models forecast new infections and deaths at the international, national, and state levels. The Centers for Disease Control and Prevention (CDC) regularly compiles on its website the forecasts of nine COVID-19 models, including LANL's. Some show the rate of new deaths slowing nationally; others show daily fatality numbers remaining flat. Most of the included models assume the continuance of the social distancing policies that were in place on the date of model calibration. A few make no such assumptions.

The unknowns about the disease and its transmission produce large error bars in predictions. The LANL model, for example, estimates that by 10 June Georgia will have 40 000 total confirmed cases of COVID-19. But that number lies between a best-case scenario of 33 000 cases and a worst case of 64 000. The lab's earlier predictions for Georgia's COVID-19 case and death totals have been closer to the best-case scenario. Osthus says the model's skill varies widely from state to state, depending in large part on the quality of data used as input. It has performed well for New Mexico, for example, but poorly for Connecticut. The model has a built-in assumption that the growth rates of new cases and deaths will trend toward zero, but it is constructed so that it can adapt to new and unexpected input, he says.

Statistics and mechanics

At the most basic level, models are either statistical or mechanistic attempts to simulate the disease transmission cycle. Until early May, the IHME model was purely statistical; no assumptions were made about how the disease spreads. Instead, it assumed that the epidemic will follow a trend that is based on experiences in China, South Korea, Italy, and other places.

On 4 May, after many states had begun reopening businesses, IHME updated its model to include a disease-transmission component, which took into consideration the relaxation of states' stay-at-home orders. The hybrid model produced more somber conclusions, including a near doubling of the US death toll to 135 000 by 4 August, up from the 72 000 deaths the old model had predicted. As of 13 May, US fatalities from the new coronavirus stood at nearly 83 000, according to the authoritative COVID-19 tracker maintained by Johns Hopkins University.

A statistical model can be appealing, says Harvard University research fellow Alison Hill, "because it fits the data well. It describes what really happened, and that maybe it's reasonable to expect that the trend will keep going in a particular direction." Such models have been used to compare the anticipated number of COVID-19 cases to the regional capacity of hospitals.

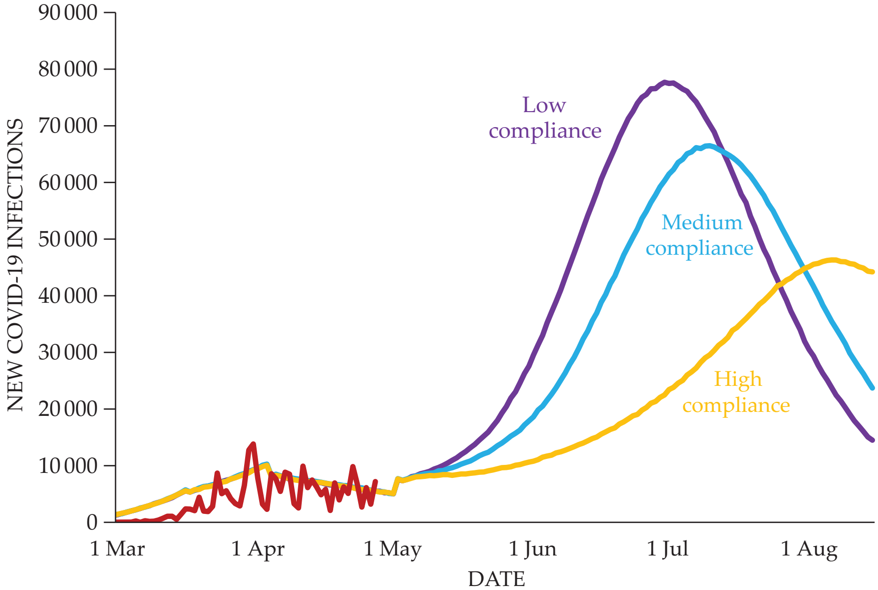
Purely statistical models, however, have plenty of limitations. "There's no natural law that says epidemics have to follow that curve," says Hill. Statistical models can't predict the outcome of relaxing restrictions and a resurgence of the disease, nor can they be adjusted for varying degrees of adherence to social distancing measures. There's no clear way to adapt a statistical model to accommodate multiple disease peaks or to consider what a combination of widespread testing and contact tracing could do to alter the pandemic's course, she says.

Mechanistic modeling, on the other hand, seeks to emulate the disease transmission process. MIT's model, developed by a team led by business professor Dimitris Bertsimas, is one example of a commonly used mechanistic approach known as SEIR, named for the successive phases—susceptible, exposed, infected, and resistant—through which it tracks the progression of the disease. The IHME's sharp increase in mortality predictions reflected the addition of a new SEIR component to its model's statistical underpinnings.

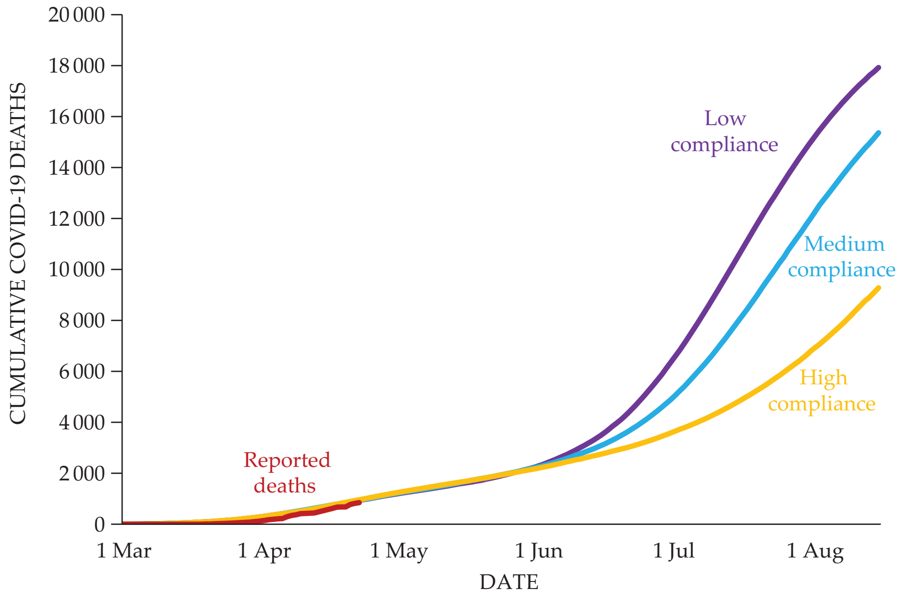
Mechanistic models are akin to physical models: Both translate laws and processes into mathematical form. Philip Stamp is a condensed-matter theorist at the University of British Columbia. He was called in by Brett Finlay, a UBC microbiologist who was first to map the genome of the 2003 severe acute respiratory syndrome (SARS) coronavirus, to formulate equations for a COVID-19 forecasting model being developed for the provincial government. "This is the sort of thing that theoretical physicists do, whether they are working in statistical mechanics or theoretical astrophysics," Stamp says. "Whereas for most people in medical science, it's very far away from their expertise."

Stamp says that many disease models initially failed to account for the large numbers of asymptomatic carriers of the new coronavirus. "A lot of the assumptions that went into those models were obviously wrong and could be seen to be wrong way back at the beginning of February," he asserts. From the way the disease progressed, initially in China and then elsewhere, "it was absolutely obvious that most carriers were asymptomatic. There was no other way to explain the way this evolved."

The Georgia Tech model is mechanistic and agent-based. Each of the model's 1 million agents is assigned an age, census tract, household, and peer group based on the state's demographic profile. At the start of a simulation, the infection is introduced randomly to agents according to the distribution of Georgia's confirmed cases. For infected agents, the disease progresses through its stages with predefined probabilities, and the infection spreads to healthy individuals at rates that depend on each agent's social contacts at home, in peer groups, and in the broader community. The model allows those interaction rates to be adjusted according to the degree of social distancing the agent follows.



PROGRESSION OF COVID-19 INFECTIONS in Georgia, as simulated by researchers at Georgia Tech. Shown are new daily infections for three scenarios. The simulations begin on 18 February with increasing voluntary quarantining by infected individuals and households, followed by school closures on 16 March, sheltering in place beginning 3 April, and then different levels of voluntary quarantine. The actual rate of new infections (red line) is multiplied by 8 to reflect underreporting. Other simulations (not shown) varied the shelter-in-place duration. (Adapted from P. Keskinocak et al., medRxiv preprint, doi:10.1101/2020.04.29.20084764.)



CUMULATIVE COVID-19 DEATHS in Georgia, as simulated by researchers at Georgia Tech. The simulations begin on 18 February with increasing voluntary quarantining by infected individuals and households, followed by school closures on 16 March, sheltering in place beginning 3 April, and then different levels of voluntary quarantine. The red line indicates confirmed numbers of deaths through 20 April. (Adapted from P. Keskinocak et al., medRxiv preprint, doi:10.1101/2020.04.29.20084764.)

Agent-based models are particularly useful for predicting the geographical course of the disease. Julie Swann, an engineering professor at North Carolina State University, is adapting for that state an agent-based model that she helped build at Georgia Tech for forecasting the 2009 H1N1 influenza pandemic in Georgia. The effects of social distancing, voluntary quarantining, and school closures are being built in. Data on household sizes, ages, and commuting flows are taken from US Census information.

All models make assumptions based on many unknowns about the novel coronavirus. Among the most important is the number of underreported infections. "Every respectable model that I've seen is taking into account the fact that there are plenty of people out there who are infected and haven't been reported," says Hill. The Georgia Tech model assumes that the real number of infections is six to eight times that of confirmed cases. Keskinocak says some reports suggest that actual cases could be 10 times the confirmed number.

Another important unknown is whether recovered individuals will be immune to reinfection, and for how long that immunity lasts. In the Georgia Tech model, permanent immunity is assumed because there isn't definitive evidence to the contrary, Keskinocak says.

Most models, including Georgia Tech's, address what happens several weeks ahead. Projecting the track of COVID-19 in the months and years beyond the initial wave of the pandemic was the objective of researchers led by Harvard's Marc Lipsitch and Yonatan Grad. Their model draws from observed experience with coronaviruses that cause the common cold and tries to estimate some of the epidemiological parameters for COVID-19 that will determine its long-term dynamics. Depending on factors such as the duration of immunity in recovered patients, the degree to which warmer weather causes COVID-19 to recede, and the extent to which immunity to cold viruses might extend to COVID-19, the model presents multiple scenarios for potential outbreaks through 2025. Absent effective vaccines or therapeutics, aggressive contact tracing, quarantining, and intermittent social distancing may need to be maintained into 2022, the authors conclude, at a substantial social and economic cost.

Decision-making tools

"All forecasts are wrong, because you are predicting the future," says Swann. "But if you know how to interpret them, you can come to understand what aspects to look at, and you can offer some real benefits for decision making." If used correctly, a model will indicate what to expect in terms of infections and deaths weeks after schools are reopened, she says. Policymakers will have to consider the modeling results in the context of other costs—unemployment, poverty, suicides, and so on—resulting from continuing stay-at-home orders. Such tradeoffs are made all the time by governors, legislators, mayors, and city council members, Swann notes.

"Models are absolutely crucial in the same way that economic models help in planning the economy," says Stamp. "But for the virus, the set of equations that tell you how this thing will evolve have not been there. [Virus modelers] have been using some simple equations that aren't working, and they're simply not an adequate description of what's out there."

With so many models to choose from, which should policymakers rely on to guide them? The CDC compiles an ensemble of the nine models that it tracks. It generally plots a middle course between the most dire and most hopeful predictions.

"It can be a good thing to use more than one model, because you don't want to be fully dependent on a set of assumptions without really challenging those assumptions," says Swann.

"This is just way too hard of a problem to think you are going to get it right, but it's also incredibly consequential," says Osthus. Given the model's varying success from state to state, LANL began including performance records in each state forecast. "Decision makers are looking at this model and trying to base meaningful decisions on it, and we think that they are entitled to this information," he says. "The user can look at our model performance and decide if they want to trust what it's saying."

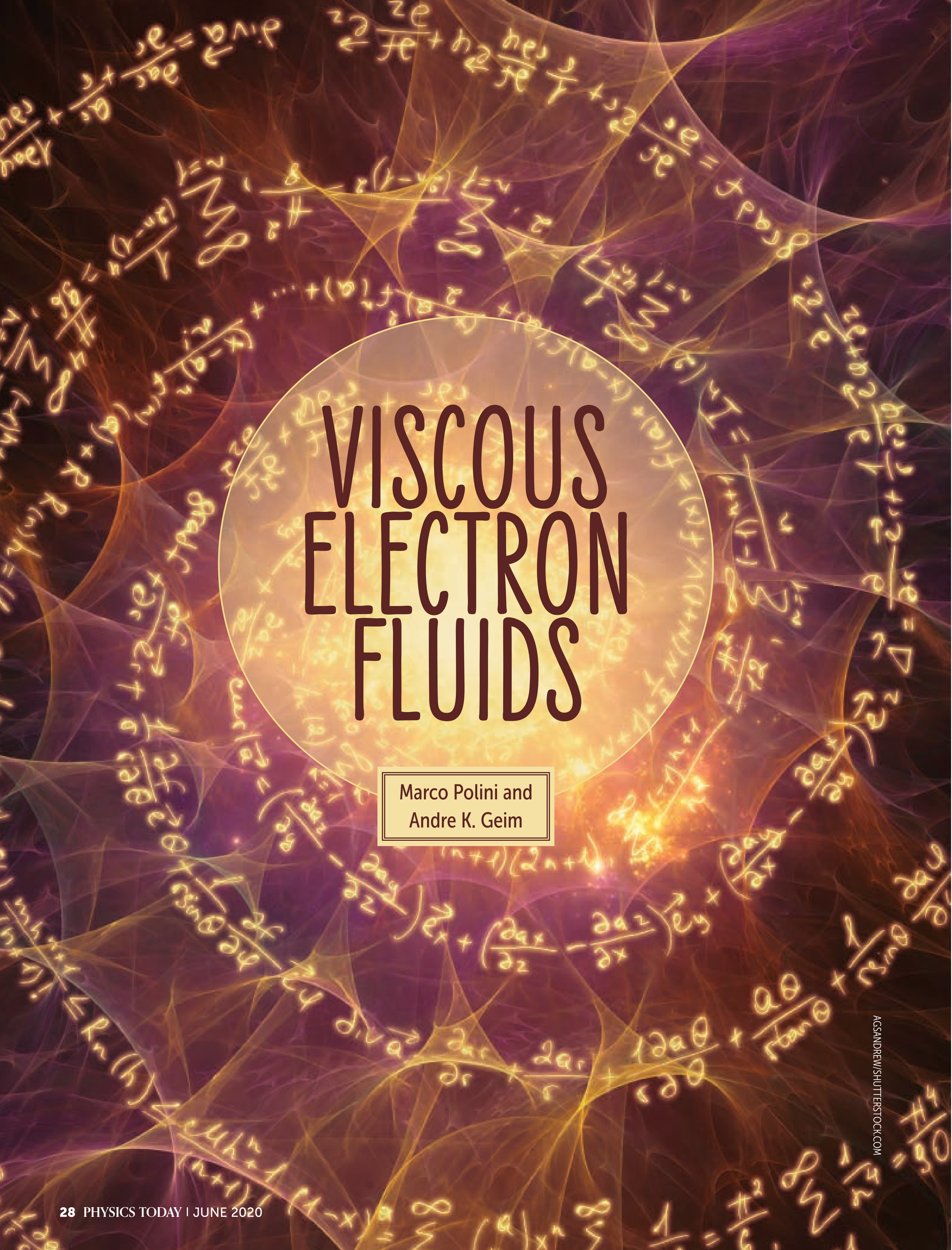
One factor to look for in evaluating models is the training and experience of the people who make them, says Hill. "There are always people trying to jump in and give their two cents' worth on what they think will happen."

There's little expectation that a single, authoritative model will emerge. Hill says researchers will never agree on what assumptions should be given more weight relative to others. "People who make models and communicate them need to be more straightforward in explaining in an accessible way what assumptions they've made," she says. "We'll never get away from uncertainty and think that people will agree on what's certain and what's not."

David Kramer PT

VISCOUS ELECTRON FLUIDS

Marco Polini and
Andre K. Geim



AGSANDREW/SHUTTERSTOCK.COM

Marco Polini is a professor at the University of Pisa in Italy, a professor at the University of Manchester in the UK, and an external collaborator of the Italian Institute of Technology in Genoa. **Andre Geim** is the Regius Professor at the University of Manchester.

Advances in materials science have made it possible for electrons in metals to exhibit exotic hydrodynamic effects.

Electrons in metals and semiconductors are often naively described as little balls bouncing around, much like atoms or molecules in dilute gases. That description, sketched in figure 1, originally came from Lev Landau, who reduced the complex many-body problem to a Fermi gas of nearly free electrons. But his simplification is counterintuitive, because Landau theory also infers that electron gases in normal metals should be exceedingly viscous because of pervasive electron–electron (e–e) collisions in solids. Indeed, the theory predicts that viscosity becomes infinite with decreasing temperature T, and simple estimates show that as T drops to that of liquid helium, electron gases in metals should be more viscous than honey.

One might expect researchers to rely on hydrodynamics, with its difficult-to-solve Navier–Stokes equation, to describe the resistivity of metals. And yet that approach is not routine. The nearly free electron model works well because the spatial interactions between electrons in metals differ so much from those between molecules in a gas. Whereas molecules scatter only when they directly contact each other, the mean free path, l_{ee}, at which electrons effectively scatter is much longer—typically microns at liquid-helium T—and it grows longer still at lower T.

The large mean free path leaves plenty of time and space for impurities and thermal vibrations (phonons) to destroy any nascent collective response of electrons that would otherwise produce viscous flow. To understand why the destruction happens, imagine a highly viscous classical gas moving through a large tube, its flow experiencing dissipation, convective nonlinearities, and other hydrodynamic behavior. Now fill the tube with sand, so that the intergranular gaps are smaller than the molecules' mean free path. The flow through the porous sand would then no longer be viscous. Rather, it would be diffusive, each particle moving independently of the others.

Something similar happens in normal metals: Impurities and phonons act like those grains of sand, which are packed densely enough to eliminate any sign of the electrons' collective behavior. In theory, it should be possible to recover the intrinsic hydrodynamic behavior of electrons, unmasked by impurities or phonons, if a metal is ultraclean and cooled to a low enough T to avoid phonon scattering. (Think of that recovery as equivalent to the removal of sand.) But in practice, little experimental progress has been made in reaching that hydrodynamic regime, despite efforts over many decades. Fortunately, the availability of new high-quality electronic materials—graphene, in particular—has recently improved the situation.

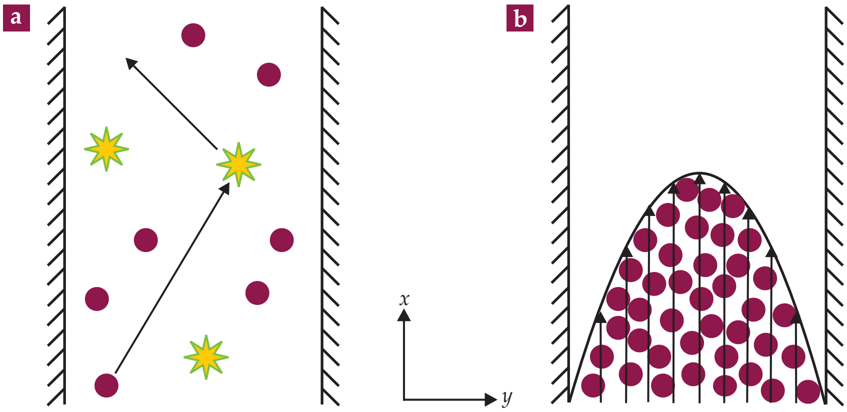


FIGURE 1. DIFFUSIVE TRANSPORT VERSUS VISCOUS-ELECTRON FLOW. (a) In the single-particle, diffusive model, electrons (red circles) move as independent particles, undergoing collisions with impurities, phonons (yellow stars), and boundaries. **(b)** In the hydrodynamic regime, in which the electron–electron mean free path is the shortest length scale in a material, frequent interactions among electrons can give rise to collective, viscous behavior, dubbed Poiseuille flow. The velocity in the flow direction, v_x, is a parabolic function of the transverse coordinate y. (Images from Marco Polini.)

A history of misbehaving electrons

In 1963 Soviet theorist Radii Gurzhi asked how a viscous electron flow could reveal itself in an experiment.[1] He assumed the existence of a metallic system in which l_{ee} was the shortest length scale that electrons would travel, much shorter than both the sample size W and the mean free path l of electrons whose collisions—for instance, with phonons and crystal defects—did not conserve momentum. Given that assumption, frequent collisions between electrons should be able to establish a collective flow, illustrated in figure 1b, because their total momentum and energy is not lost to the outside world.

Gurzhi found that the resistance R of such an imaginary metal would have to decrease with increasing T. That's a shocking result because the standard definition of a metal is that its R increases with T. Nonetheless, the theoretical prediction was unambiguous and could be traced to the fact that the electron viscosity ν in metals decreases with T. Intuitively, it also makes sense: As a system warms, it becomes less viscous, which allows easier passage of a fluid. The anomaly is usually referred to as the Gurzhi effect, and it explains that if a metal enters the hydrodynamic regime—where $l_{ee} \ll W$ and l—it should exhibit a T dependence that is the opposite of metallic and is more like that of semiconductors, whose R decreases with T.

Unfortunately, finding a system that satisfies those conditions turned out to be nearly impossible. One usually thinks of large, clean crystals at cryogenic T, which would mitigate the effect of phonons and thus increase l. Indeed, clean, three-dimensional metals at low T exhibit values of l that are nearly a centimeter. However, l_{ee} also rapidly increases with decreasing T because of what's known as Pauli blocking. (See box 1 for details on the fundamental properties of electron systems.)

Fermi statistics greatly limits the available phase space for e–e collisions when T is well below the Fermi temperature T_F. As a result, l_{ee} diverges as $(T_F/T)^2$ with decreasing T. That low-T regime is precisely where Landau quasiparticles are long lived and the single-particle model of electrical conductivity is justified.

The only way to reach the hydrodynamic regime is to work at elevated T, such that the Fermi sphere becomes "softer" and

BOX 1. A PRIMER ON FERMI LIQUIDS

Statistical mechanics tells us that the ground state of a system of noninteracting electrons is a Fermi sphere—that is, all the states with wavenumber $|\mathbf{k}|$ smaller than a maximum, dubbed the Fermi wave number k_F, are occupied, and states with $|\mathbf{k}| > k_F$ are empty. The occupation number n_k of a state with momentum $\hbar\mathbf{k}$ is therefore a step function, changing discontinuously from 1 to 0 as $|\mathbf{k}|$ crosses k_F. The energy of the state at k_F is the Fermi energy E_F, and the related temperature scale $T_F = E_F/k_B$ is the Fermi temperature. At finite T, such a step is smeared around $|\mathbf{k}| = k_F$ into a smooth Fermi–Dirac distribution function.

In a series of brilliant papers in 1957, Lev Landau showed that when electron–electron interactions are taken into account, they do not modify that single-particle picture much. In a Fermi liquid at $T = 0$, n_k displays a finite jump in amplitude when $|\mathbf{k}|$ crosses k_F. Due to electron–electron interactions, bare electrons become "dressed" electrons, known as quasiparticles. In a Fermi liquid, scattering between quasiparticles is heavily constrained by the Pauli exclusion principle; transitions can only occur between initial occupied states and final empty states. At finite T, only partially occupied states in a window of width k_BT around E_F can participate in the scattering. That "Pauli blocking" is at the heart of the existence of Fermi liquids and is responsible for the $1/T^2$ divergence in the mean free path of electrons in the limit of $T \ll T_F$.

Pauli blocking less obstructive to e–e scattering. At those higher T, phonons become the main hindrance and limit l to the electron–phonon scattering length, l_{ep}. The resulting condition, $l_{ee} \ll l_{ep}$, required to observe viscous behavior is extremely difficult to satisfy because l_{ep} often decreases faster with increasing T than does l_{ee}. (For 3D metals, l_{ep} usually varies as T^{-3}, whereas l_{ee} varies as T^{-2}.) That scaling narrows the materials systems one could use and the T interval in which electron hydrodynamics could possibly be observed.

An elegant attempt to break the impasse[2] was undertaken in the 1990s. Researchers applied a high electrical current that increased the electron T of a 2D electron semiconductor (2DES) system and shortened l_{ee}. Even so, the crystal lattice remained close to liquid-helium T, which kept the electron–phonon scattering low as well. Measuring the differential resistance revealed a small but distinct bump as a function of applied current, a feature the researchers interpreted as plausible evidence for the Gurzhi effect.

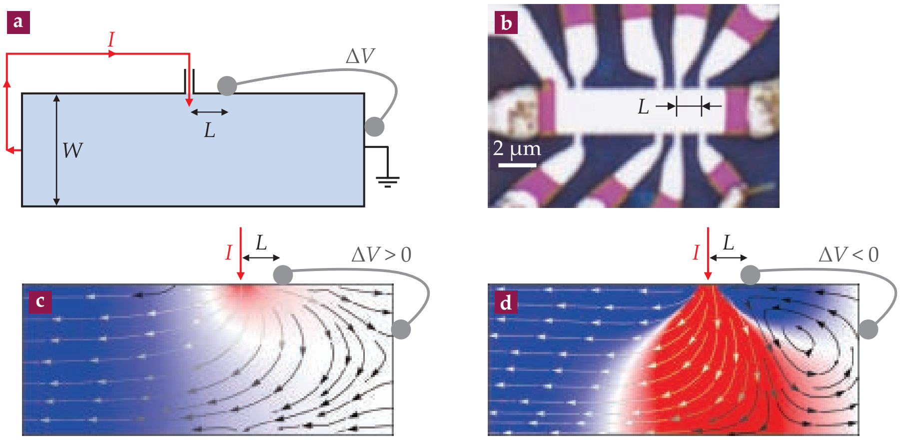


FIGURE 2. NEGATIVE RESISTANCE AND CURRENT WHIRLPOOLS. (a) In a so-called vicinity-resistance measurement, current I is injected into a two-dimensional device of width W through a narrow lead, and a potential drop ΔV is measured between a voltage probe placed a short distance L from the injector and a faraway lead. **(b)** In this micrograph of a real device, graphene (white) is tipped with electrical contacts (magenta), and current and voltage probes can sample any of several positions during an experiment. **(c)** This color map shows the calculated distribution of electrical potential in the absence of viscosity. The voltage and resistance are positive (red) and arrows reveal the steady-state current pattern. **(d)** In the case of viscous flow, lobes of negative voltage (blue), and thus negative resistance, emerge near the current injector I. The finite viscosity induces whirlpools in the current flow. (Adapted from ref. 4.)

Gurzhi and coworkers immediately disagreed with that interpretation,[3] and they pointed out that peculiarities of e–e scattering in 2D materials demand an even more stringent condition than that in 3D metals—namely, $l_{ee} \ll W(T/T_F)$, which had not been achieved in the experiment. Their rejection left the research status in limbo: For a half century after the Gurzhi theory was postulated, no electronic system had been found to exhibit unambiguous signs of hydrodynamic behavior.

Graphene to the rescue

Despite having a Nobel Prize behind it, graphene did not initially look like a promising candidate for studies of electron hydrodynamics. It was filled with impurities, with a mean free path barely exceeding 100 nm (see the article by Andre Geim and Allan MacDonald, PHYSICS TODAY, August 2007, page 35, and PHYSICS TODAY, December 2010, page 14). But that changed around 2011, when researchers found that encapsulating graphene in hexagonal boron nitride dramatically improved its electronic quality. The encapsulation shielded graphene from outside impurities and flattened the crystal by suppressing scattering at microscopic corrugations.

Today, graphene is one of the highest quality electronic materials ever produced: Its low-T mean free path is currently limited only by the device size W, at least up to 10 μm, and exceeds a micron even at room T. More importantly, graphene is extremely stiff, a feature that suppresses phonon scattering and increases l_{ep}. And unlike what happens in 3D metals, electron–phonon scattering in 2D graphene increases slowly with T; $l_{ep} \propto T^{-1}$, with a small proportionality coefficient that accounts for stiffness. As noted earlier, e–e scattering rises much faster, with $l_{ee} \propto T^{-2}$.

Therefore, above a certain T, l_{ee} is expected to become the shortest scattering length in graphene. Moreover, graphene's T_F is typically greater than 1000 K. That's neither too small, as it would be in semiconductor 2DESs, where the Fermi surface is largely destroyed at room T, nor too high for the required

BOX 2. THE NAVIER–STOKES EQUATION IN CONDENSED MATTER

The motion of water in oceans, turbulent air currents, and Marangoni flows, which produce "tears of wine" inside a glass, are a few examples of phenomena governed by the Navier–Stokes equation. The equation is essentially Newton's second law for each fluid element—a small volume of a liquid or gas subjected to external forces. Today, no mathematical theory exists that would unlock the equation's complete solution. Finding it remains one of the famous seven Millennium Prize problems.

To describe a steady-state flow of electrons, the simplest, linearized form of the Navier–Stokes equation is normally used:[4–6]

$$\frac{\sigma_0}{e}\nabla\phi(\mathbf{r}) + D_\nu^2\nabla^2\mathbf{J}(\mathbf{r}) - \mathbf{J}(\mathbf{r}) = 0,$$

in which $\mathbf{J}(\mathbf{r}) = n\mathbf{v}(\mathbf{r})$ is the current density, n is the electron density, $\phi(\mathbf{r})$ is the electric potential, σ_0 is the diffusive conductivity, and e is the electron charge. The length over which the flow's momentum diffuses is given by $D_\nu = \sqrt{\nu\tau}$, where τ is a time scale that describes momentum dissipation from the scattering of electrons with impurities and phonons. In the limit where D_ν goes to 0, the linearized Navier–Stokes equation yields Ohm's law locally: $-e\mathbf{J}(\mathbf{r}) = \sigma_0\mathbf{E}(\mathbf{r})$, where $\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$ is the electric field. To find an electron flow pattern, the Navier–Stokes equation needs to be solved together with the continuity equation, $\nabla \cdot \mathbf{J}(\mathbf{r}) = 0$, and the boundary conditions.

2D condition $l_{ee} \ll W(T/T_F)$ to be consequential. In short, it is hardly possible to imagine a better material than graphene for studying viscous electron flows.

Despite the promise of that expectation, in 2015 when the first contemporary experiments began probing the phenomenon, graphene's resistance showed no sign of the Gurzhi effect at any T. In hindsight, one can understand why the viscous effects did not show up straightforwardly. The kinematic viscosity ν enters the Navier–Stokes equation as a coefficient in front of the second spatial derivative of velocity $v(x,y)$ (see box 2). In the standard resistance measurements that use a long strip of a uniform width, only $v_x(y)$—the y dependence of flow velocity in the x direction—is nonzero. Unless significant momentum losses occur at the strip boundaries, the y dependence tends to be weak. The result is a fairly uniform flow profile. And without a significant velocity gradient, the viscosity term contributes little to the solution of the Navier–Stokes equation and, hence, to the resistance R.

That insight offered a tip for how to proceed: To maximize the hydrodynamics effects in experiment, it is essential to create a current flow as inhomogeneous as possible.[4]

Negative resistance and whirlpools of electrical current

One geometry that provides large velocity gradients is a narrow current injector, shown schematically in figure 2. According to the Navier–Stokes equation, the electric potential changes its sign at a characteristic distance of order of $D_\nu = \sqrt{l_{ee} l}/2$ from the injector.[4–6] One can measure that local potential in the so-called vicinity geometry—that is, by placing a voltage probe sufficiently close to the injector. The corresponding resistance R_V—the local voltage divided by the injected current—has the normal, positive sign for noninteracting electrons in both diffusive and ballistic transport regimes. Negative R_V, by contrast, is a smoking gun for viscous flow.[4]

However, one must be careful. As T increases, the initial sign change indicates that ballistic transport is strongly affected by e–e interactions, and the hydrodynamic regime develops only later, at higher T when collisions among electrons become more frequent.[7] The observation of negative R_V in graphene and its comparison with behavior expected by Navier–Stokes theory allowed the first measurements of an electron fluid's viscosity. At liquid-nitrogen T, ν turns out to be 100 times as great as honey. Reassuringly, that result agrees with many-body theory.[4]

Navier–Stokes theory also predicts another spectacular effect in the conductivity of metals because of viscosity.[4–6] The negative region of electric potential near the injector is predicted to develop into a whirlpool of electrical current. Whirlpools are familiar phenomena in the laminar flow of ordinary fluids, but in the vicinity geometry[4,6] of figure 2a, they are theoretically expected to exist near a narrow injector. Only the size of D_ν depends on the actual value of ν.

For other geometries that create a nonuniform flow,[5] current whirlpools generally disappear if D_ν gets smaller than the characteristic device size W, even though the negative potential anomaly doesn't change.

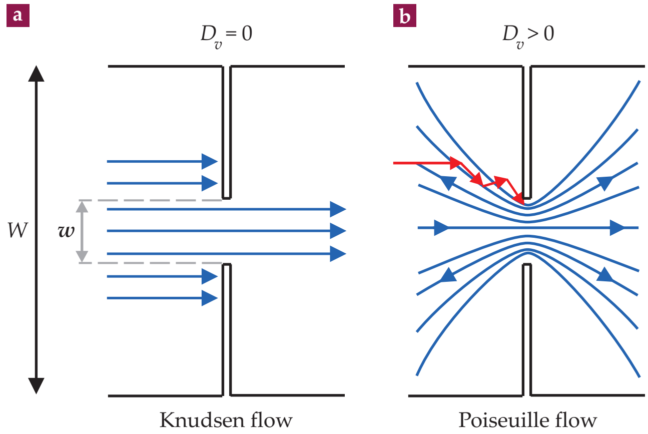


FIGURE 3. ELECTRON FLOW THROUGH A CONSTRICTION. A narrow aperture of width $w \ll W$ separates two wide leads. **(a)** In ballistic transport—Knudsen flow in the language of gas dynamics—electrons move independently. With no scattering between them, the resistance to their flow (blue) through the constriction had been expected to be a minimum. **(b)** In a viscous electron fluid, however, Poiseuille flow corresponds to yet lower resistance. An individual electron (red), initially directed toward the boundary, isn't expected to contribute to the conductance. But collisions with other electrons effectively drag it toward the constriction and the collective motion decreases the resistance. The quantity D_ν is the length scale over which momentum diffuses as a result of electron–electron collisions. (Images from Marco Polini.)

Electrons go superballistic

In 1908 Martin Knudsen observed that the speed of gas flowing through a small aperture suddenly increased when he increased

BOX 3. ANTI-MATTHIESSEN'S RULE

Formulated in 1864, Matthiessen's rule states that if several independent scattering processes exist in a system, the total resistance R is the sum of the resistances due to each process. Deviations from the rule occur in metals but are generally tiny. The occurrence of an anti-Matthiessen's rule, in which conductivities G rather than resistivities are added, is exceptionally rare. One possible scenario was proposed for the case of strange metals.[14–16]

A viscous electron flow through a point contact (PC) is another exception. Two relevant time scales exist in that situation. The first is the single-particle flight time across the constriction, $\tau_1 = 2/\pi(w/v_F)$, where w is the size of a constricting aperture and v_F is the Fermi velocity. The second is the time scale over which the momentum diffuses over the same distance, $\tau_2 = \pi/32(w^2/\nu)$, where ν is the viscosity. The total PC resistance[9] is given by

$$R_{PC} = \frac{1}{G_1 + G_2} = \frac{m}{n e^2} \frac{1}{\tau_1 + \tau_2},$$

where m is the effective electron mass, G_1 is the Sharvin conductance, G_2 is the contribution to conductance from electron–electron interactions, and n is the electron concentration. Three years ago, experiments confirmed the validity of that anti-Matthiessen equation.[8]

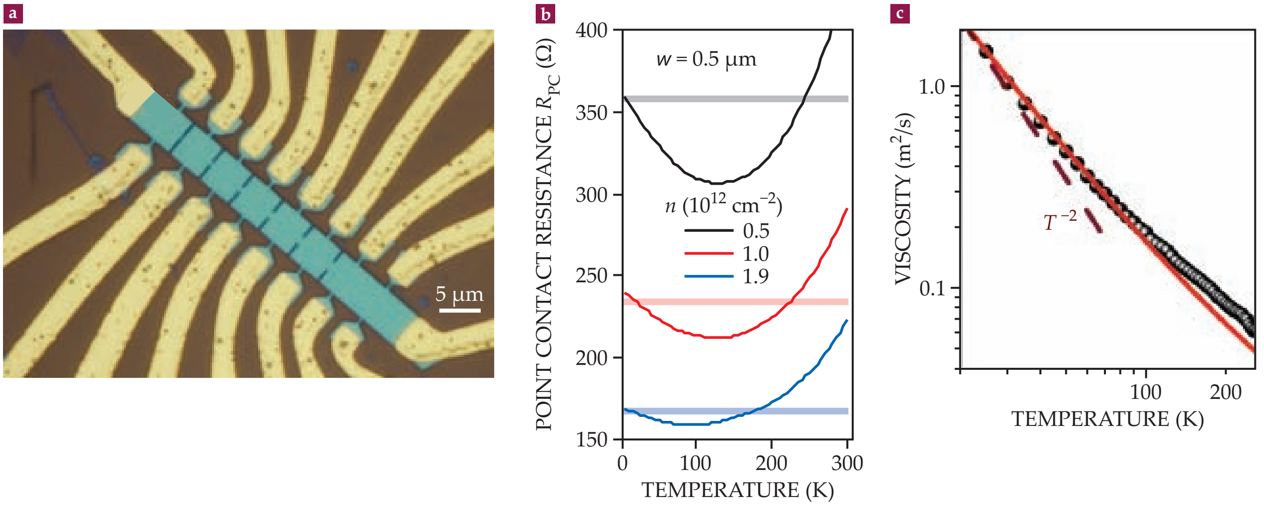


FIGURE 4. THE GURZHI EFFECT. (a) A graphene device has a series of point contacts of different widths w; the contacts link several boxes (turquoise), each connected to separate electrodes (yellow). **(b)** The resistance of one of those point contacts (w = 0.5 μm) is plotted as a function of temperature T for three electron densities n. The horizontal lines indicate the ideal, ballistic limit. But as T increases, the resistance drops below the expected minimum and follows a nonmonotonic dependence on T—the Gurzhi effect. **(c)** Black dots represent viscosity[8] measured as a function of T for $n = 10^{12}$ cm^{-2}. The experimental dots closely agree with many-body theory calculations (red line). For comparison, note the y-axis scale: The viscosity of honey is about 10^{-3} m²/s. (Adapted from ref. 8.)

the gas's density. The experiment implies that a higher viscosity boosts the gas flow, which is a counterintuitive result. The effect is well understood today as the transition from Knudsen flow to Poiseuille flow, or in the language of metal physics, from ballistic-electron transport to viscous-electron transport. The phenomenon observed by Knudsen can be viewed as the analogue of the Gurzhi effect for gases rather than electrons.

An experiment similar to Knudsen's was recently performed on graphene.[8] As shown in figures 3 and 4, a narrow aperture of width w connects two wider regions, a geometry known as point contact (PC). In the ballistic regime at low T, such PCs were first made and studied by Yuri Sharvin in the 1960s. He found that even in the ideal case—without any disorder and scattering—a PC exhibited finite electrical conductance. Its value is given by the number of electron-wave modes that can fit inside the aperture.

Until recently, researchers have tacitly accepted that Sharvin's conductance was the highest possible value. The absence of disorder seemed to imply the best-case scenario for unimpeded electron transport. But that turned out to be wrong. Figure 4b shows that when T is increased and a system enters the hydrodynamic regime, the resistance measured in a graphene PC drops below the ideal ballistic limit. For the experiment in the figure, the drop was caused by the transition from ballistic to viscous electron transport. It was also accompanied by a semiconductor-like T dependence—the first unambiguous manifestation of the Gurzhi effect.

How is it possible for viscosity to lower the electrical conductivity? After all, basic physics tells us that greater electron scattering should increase the resistance—a trend known as Matthiessen's rule. Making the transition from the low-T regime, where Sharvin's description applies, to the higher-T hydrodynamic regime, electron viscosity sets up a funnel-like current pattern through the aperture, akin to what happened in Knudsen's experiment.

Imagine an electron moving toward the PC, as in figure 3. In the ballistic regime, it hits the wall and stops contributing to the conductance. But in the hydrodynamic regime, the same electron is dragged by electron collisions toward the opening and forced to funnel through it. That funneling is what raises the conductance above Sharvin's ballistic limit. Mathematically, the superballistic flow happens because conductivities are added—the so-called anti-Matthiessen's rule[9] described in box 3. By comparing experimental results and theory, the two of us and our colleagues were able to accurately measure graphene's viscosity as a function of electron concentration and T.

Electronic magnetohydrodynamics

Another knob that can be turned to explore viscous flow is the magnetic field **B**. In traditional metallic systems, **B** causes the Hall effect, a potential drop perpendicular to the direction of both current flow and the magnetic field. How is the Hall effect influenced by electron viscosity? The presence of a magnetic field breaks down time-reversal symmetry and produces a new kinematic coefficient ν_H in the Navier–Stokes equation. The coefficient, known as the Hall viscosity, is odd under reversal of **B** and is dissipationless. The Hall viscosity gives rise to an extra term in the Navier–Stokes equation that is proportional to ν_H, acts against the Lorentz force, and suppresses the resulting potential drop.

The suppression of the Hall effect is local and extends only over distances of D_ν, typically about 0.5–0.6 μm. By placing voltage probes close to a narrow current injector, we measured a local Hall effect.[10] For graphene in the hydrodynamic regime, it was found to be notably smaller than the standard Hall effect, measured simultaneously at some distance from the current contact.

What's next

Now that we know how to force hydrodynamics to show up in experiments, we expect to soon observe viscous phenomena in many systems, including 2DESs in semiconductors, graphite, bismuth, topological insulators, and Weyl metals. Evidence already exists for viscous flow in delafossites,[11] and local (vicinity and PC) geometries should help make those observations. Materials in which electrons and holes coexist and interact strongly present another interesting challenge.[12,13]

Let's also not forget about materials that defy the Fermi-liquid paradigm. They are called strange metals[14,15] and have Planckian transport scattering times on the order of $\hbar/(k_B T)$ down to the lowest T. Those metals are also expected to exhibit viscous electron motion, albeit with a tiny viscosity conjectured to be close to a universal lower bound predicted by string-theory methods. Experimental evidence of the lower bound has been reported in ultrahot nuclear matter, such as quark–gluon plasmas, and in ultracold atomic Fermi gases, but not in condensed-matter physics.

Yet another enticing project would be to extend existing hydrodynamic studies into the regime where nonlinear terms in the Navier–Stokes equation can no longer be ignored. In classical fluids, those terms are responsible for nonlinear phenomena such as turbulence. Similar physics is expected to occur in electron fluids, but studying such fluids would require materials with smaller ν and longer τ compared with the 2DESs studied so far.

For all those new ventures, one should use not only electrical probes but also the visualization tools that are now available. Scanning probe microscopes that can sense voltages or magnetic fields are one example. They can image local distributions of electrical current at submicron scales and reveal electron hydrodynamics at an entirely new, more spectacular level. Watch out for beautiful images of electron whirlpools and viscous flows coming soon.

The European Union's Horizon 2020 research and innovation program (Graphene Flagship) supported this work. We are grateful to everyone who contributed to the research—particularly Denis Bandurin, Alexey Berdyugin, Roshan Kumar, Leonid Levitov, Francesco Pellegrino, Leonid Ponomarenko, Alessandro Principi, Andrea Tomadin, and Iacopo Torre. Author Marco Polini dedicates this article to Rachele.

REFERENCES

1. R. N. Gurzhi, *J. Exp. Theor. Phys.* **17**, 521 (1963).
2. M. J. M. de Jong, L. W. Molenkamp, *Phys. Rev. B* **51**, 13389 (1995).
3. R. N. Gurzhi, A. N. Kalinenko, A. I. Kopeliovich, *Phys. Rev. Lett.* **74**, 3872 (1995).
4. D. A. Bandurin et al., *Science* **351**, 1055 (2016).
5. L. Levitov, G. Falkovich, *Nat. Phys.* **12**, 672 (2016).
6. F. M. D. Pellegrino et al., *Phys. Rev. B* **94**, 155414 (2016).
7. D. A. Bandurin et al., *Nat. Commun.* **9**, 4533 (2018).
8. R. K. Kumar et al., *Nat. Phys.* **13**, 1182 (2017).
9. H. Guo et al., *Proc. Natl. Acad. Sci. USA* **114**, 3068 (2017).
10. A. I. Berdyugin et al., *Science* **364**, 162 (2019).
11. P. J. W. Moll et al., *Science* **351**, 1061 (2016).
12. J. Crossno et al., *Science* **351**, 1058 (2016).
13. P. Gallagher et al., *Science* **364**, 158 (2019).
14. For a popular introduction to strange metals, see J. Zaanen, *SciPost Phys.* **6**, 061 (2019).
15. S. A. Hartnoll, A. Lucas, S. Sachdev, *Holographic Quantum Matter*, MIT Press (2018).
16. P. A. Casey, P. W. Anderson, *Phys. Rev. Lett.* **106**, 097002 (2011). PT

A cylindrical detector in the LADI instrument at the Institut Laue–Langevin in Grenoble, France, records neutron diffraction patterns. The data enable the generation of three-dimensional molecular structures that include hydrogen atoms. (Bob Cubitt/ILL.)

David Hoogerheide is a researcher at the National Institute of Standards and Technology Center for Neutron Research in Gaithersburg, Maryland. **Trevor Forsyth** is a senior fellow and head of the life sciences group at the Institut Laue–Langevin in Grenoble, France; he also holds a chair in biophysics at Keele University in the UK. **Katherine Brown** is a senior researcher in the Cavendish Laboratory at Cambridge University in the UK and at the University of Texas at Austin.

David P. Hoogerheide, V. Trevor Forsyth, and Katherine A. Brown

Neutron scattering for STRUCTURAL BIOLOGY

Modern neutron sources illuminate the complex functions of living systems.

The past two decades have seen explosive growth in research on structural molecular biology. High-throughput techniques for determining biological structures are yielding large amounts of information about atomic-, cellular-, and tissue-scale organization. Advances are driven by modern high-brilliance synchrotron sources, synchrotron-based full-field x-ray microscopy and tomography, free-electron lasers, and cryoelectron microscopy facilities. The scientific landscape is changing at a remarkable pace with increasing emphasis being placed on interdisciplinary and multi-technique approaches. Neutron scattering facilities around the globe are expanding their capabilities to provide unique and complementary insights about biological systems.[1]

Neutron-based techniques have become increasingly powerful and sophisticated tools for structural biology research. Recent improvements in sources, instrumentation, and sample preparation have revolutionized the scope of biological neutron scattering. Although such methods have been used to probe macromolecules[2] since the 1960s, their application to the study of biological systems was limited because data acquisition was technically difficult and frustratingly slow. (For more about early work on biological neutron scattering, see the article by Peter Moore, Physics Today, January 1985, page 62.) Technological advances are changing how researchers perceive and exploit neutron scattering. Here, we focus on current and emerging techniques that are propelling cutting-edge research in structural biology at modern neutron sources.

Neutrons around the world

There are two types of modern neutron beam sources. Continuous, or steady-state, sources produce neutrons by fission, whereas pulsed sources produce neutrons by spallation—the breakup of nuclei. The two types of sources have different yet complementary properties. Continuous sources use classical optical elements such as choppers, monochromators, and analyzers to condition neutron beams and parse scattering data. Pulsed sources have an inherent time structure that derives from their use of proton accelerators; that structure allows for time-of-flight measurements and offers the ability to discriminate between particle energies. Neutrons produced by both types of sources typically have energies similar to those associated with atomic and molecular motions. To study macromolecular systems under physiologically relevant conditions, researchers use cold neutrons with wavelengths from 40 Å to 2 Å and thermal neutrons with wavelengths less than 2 Å because both types probe biologically relevant length and time scales.

Neutron facilities that undertake structural biology research exist throughout the world (see figure 1). In the US, they include

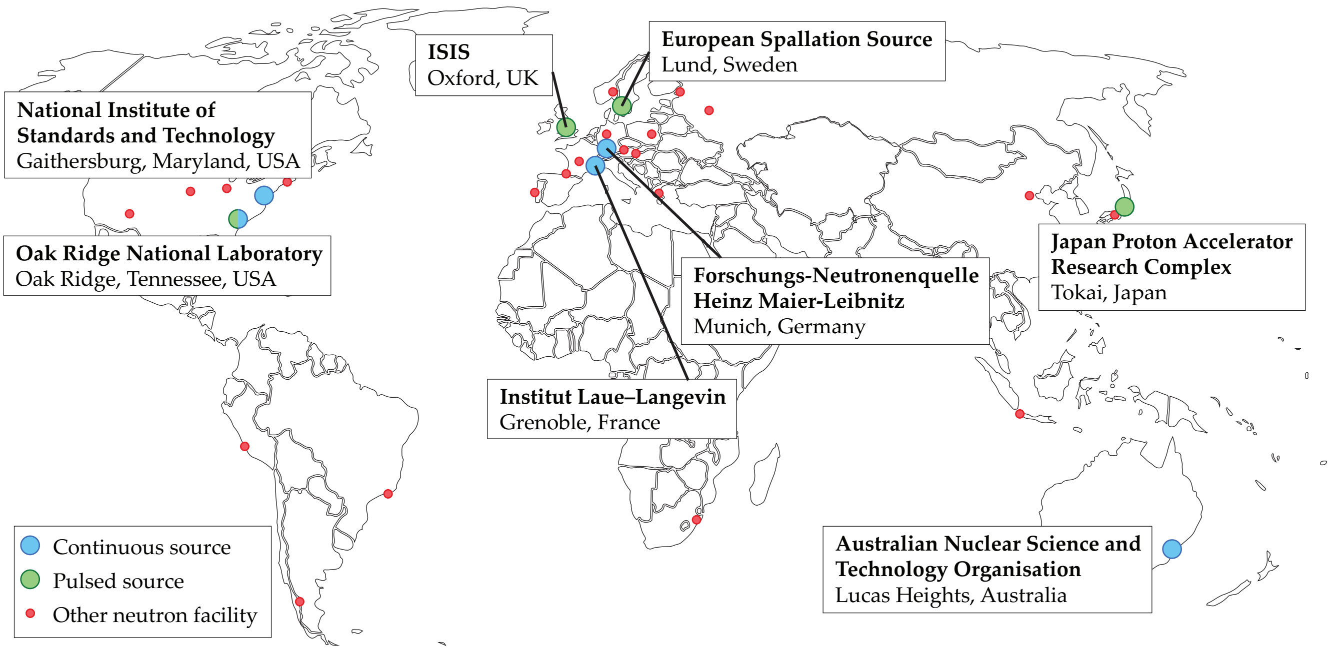


FIGURE 1. NEUTRON FACILITIES are found across the globe. Some, like the Institut Laue–Langevin, have been engaged in biological neutron research for decades. The European Spallation Source, on the other hand, is still under construction. Once operational, it will be the world's most powerful spallation source. (Tetiana Chemerys/Alamy Stock Photo; adapted by Freddie Pagani using www.neutronsources.org/neutron-centres.)

Oak Ridge National Laboratory (ORNL) and NIST in Maryland. ORNL hosts both a continuous and a spallation source, and NIST operates a continuous source. Benno Schoenborn's work on myoglobin at Brookhaven National Laboratory in the 1960s stimulated a great deal of interest in applying neutrons to the study of biological systems. Los Alamos National Laboratory has also made significant contributions to structural biology.

In Europe, the Institut Laue–Langevin (ILL) in Grenoble, France, founded in 1967, pioneered the use of neutrons to study biological systems. Groundbreaking research at the ILL by Heinrich Stuhrmann and Bernard Jacrot drove an increase in the use of neutrons for studying macromolecules in solution. It spurred the establishment of a European Molecular Biology Laboratory (EMBL) outstation at the ILL in 1975 to support the developing field. The Forschungs-Neutronenquelle Heinz Maier-Leibnitz (FRM-II) facility in Germany has also invested considerably in structural biology research. The ISIS Neutron and Muon Source in the UK has focused on the study of soft matter with biological applications. The European Spallation Source (ESS) in Sweden is scheduled to produce its first neutrons in 2020 and should reach full power in 2024; it will be the world's most powerful spallation source and include several instruments for biological research.

Structural biology initiatives for neutron-based research are also underway in the Asia–Pacific region. The Japan Proton Accelerator Research Complex (J-PARC) spallation source produces high-intensity secondary neutron beams, and the Australian Nuclear Science and Technology Organisation neutron source has a suite of experimental equipment for biological research. The China Spallation Neutron Source, which opened in 2017, is also developing a structural biology program.

Solution scattering

Small-angle neutron scattering (SANS) is the most widely used neutron-based technique for structural biology. It involves diffracting a cold neutron beam with a solution of proteins or other macromolecules, as illustrated in figure 2a. The diffracted beam creates a scattering pattern that arises from interactions between the incident neutron beam and atomic nuclei in the macromolecule and solvent. The pattern contains information that can be used to construct images of the molecules of interest.

A unique advantage of SANS studies is that they can highlight specific parts of a molecule or molecular complex through solvent contrast variation. The approach, summarized in figure 2b, is often used in conjunction with selective deuterium labeling,[3] which is carried out by the *in vivo* synthesis of proteins in deuterated media and allows components in a complex system to be distinguished at a relatively low resolution (10–20 Å). The technique exploits the distinct coherent and incoherent neutron scattering properties of hydrogen and D atoms. Hydrogen has a large incoherent scattering cross section and a negative coherent scattering length. Deuterium, on the other hand, has a negligible incoherent scattering cross section and a large, positive coherent scattering length comparable to that of most other atoms in biological molecules.

SANS is sensitive to the relative arrangement of the proteins, nucleic acids, and lipids in assemblies of biological molecules. It can reveal structural arrangements in multicomponent systems and is therefore being increasingly used alongside complementary techniques such as small-angle x-ray scattering (SAXS), NMR, electron microscopy, and mass spectrometry. The three-dimensional structure shown in figure 2c was constructed from a combination of SANS, SAXS, and NMR measurements. The complex includes an enzyme known to chemically alter RNA as part of the normal cellular life cycle,[4] and its study provided insight into how the enzyme's components assemble and how its interactions with RNA regulate the protein's chemical modification and correct folding in living cells.

Stealth nanodisk containers have been developed for use

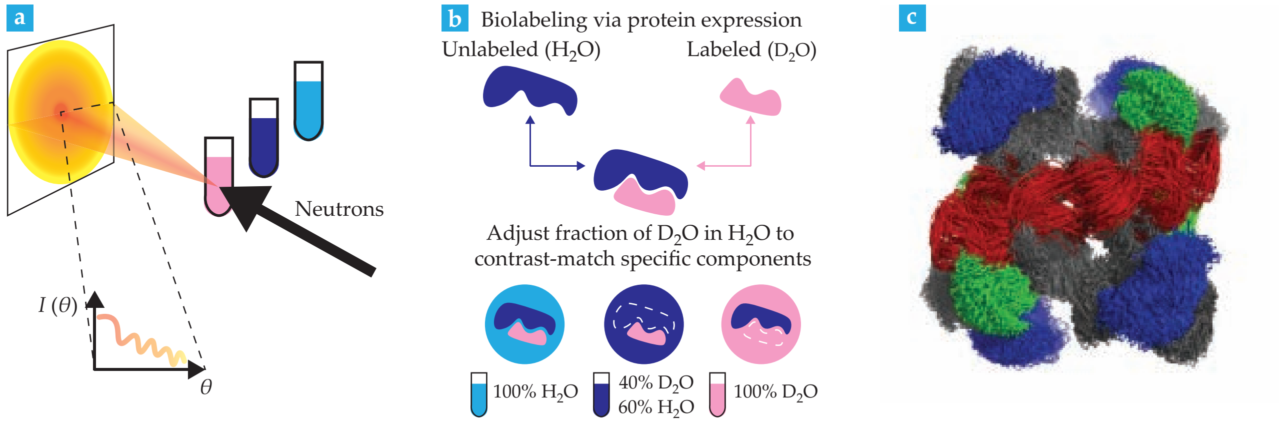


FIGURE 2. SMALL-ANGLE NEUTRON SCATTERING (SANS) probes biomacromolecular structures in solution. **(a)** In a SANS experiment, the incident beam is scattered at small angles by proteins and other macromolecules in solution. Scattered neutrons are typically collected using a two-dimensional helium ion detector, shown here as a flat screen, and the data are transformed into plots of scattering intensity I as a function of angle θ. **(b)** The scattering properties of each component in a complex can be independently varied by selectively deuterating specific parts of the structure. Adjusting the fractions of D_2O and water in the surrounding solution changes the contrast of each component.[3] **(c)** The structure of a complex formed between RNA (red) and the subunits of the box C/D ribonucleoprotein enzyme (gray, blue, and green) shown here was found using SANS in concert with other biophysical techniques. (Panels a and b courtesy of David Hoogerheide; panel c adapted from ref. 4.)

with SANS. They provide a lipid environment into which membrane proteins are inserted. Selective deuteration that renders the containers invisible enables the study of isolated membrane protein structures in a biomimetic lipidic context.[5]

More than 20 neutron sources host SANS instruments, and their capabilities for biological research are expanding on multiple fronts. For example, novel instrument and sample configurations have been developed, including integrated size exclusion chromatography and spectroscopy.[6] Anne Martel and collaborators are building an *in situ* SAXS system on the D11 instrument at the ILL. The combination of SANS and SAXS allows multiple properties of a protein's structure to be probed simultaneously. A facility having both SANS and SAXS functionalities enables the study of complex biological systems with features on scales ranging from molecular to subcellular.

Emerging computational approaches involving SANS include the development of software programs to reliably evaluate data and to apply advanced molecular dynamics (MD), Monte Carlo, and normal mode analysis methods to building biologically relevant atomistic model structures.[6,7] The Small Angle Scattering Biological Data Bank and the BIOISIS repository for SAXS data underpin those approaches by standardizing the deposition and validation of scattering data and models derived from neutron and x-ray sources.

Surface structures

Neutron reflectometry (NR) elucidates the structures of macromolecules on immobilized surfaces. Lipidic systems that mimic biological membranes are of particular interest. As shown in figure 3a, a typical NR experiment uses one or more lipid bilayers spread on a planar substrate. A temperature-controlled liquid flow cell creates a solid–liquid or liquid–air interfacial surface at which macromolecules sit. Detectors count reflected neutrons; those data are then transformed into reflectivity curves from which structural information about the interface is extracted.[8] Contrast variation and deuteration of individual components such as proteins or lipids can aid in the interpretation of NR data.

NR is useful for understanding structural arrangements and studying molecular interactions between proteins and membrane macromolecules. Like other neutron methods, it is particularly powerful when combined with other biophysical techniques. For example, NR data collected at NIST in collaboration with the National Institutes of Health were used in combination with MD simulations to understand how the protein tubulin interacts with biological membranes. Tubulin is known for its role as the major component of microtubules in the cytoskeleton, and it is a target for chemotherapeutic drugs such as taxol. However, tubulin has also been shown to associate with outer membranes of mitochondria, which are responsible for cellular respiration.

To understand how that association could occur, researchers analyzed NR data of membrane-bound tubulin using composition space modeling.[9] They described the structure of the membrane–protein complex using information such as molecular volumes, chemical connectivity, and known three-dimensional structures. Figure 3b compares the orientation of membrane-bound tubulin derived from NR with the orientation from a complementary MD simulation. Together, the two techniques revealed the membrane-binding domain and rotational dynamics of tubulin interacting with membrane lipids.

At the ILL and ISIS, NR studies have shed light on the relationship between cholesterol and lipid bilayers. By making full use of the ability to produce designer-deuterated analogues of cholesterol and deuterated lipids, the studies have improved our understanding of atherosclerosis.[10]

Crystallography

Neutron sources can determine single-crystal protein structures with atomic or near-atomic resolution. The experimental arrangement used for neutron macromolecular crystallography (NMX) is shown in figure 4a. Ideally, the sample is a crystal grown from fully deuterated protein. The crystal is maintained in a hydrated environment or cryocooled during measurements. Data are collected as Bragg reflections using image-plate, scintillator, or gas detectors. The data are then used to produce a density map, yielding an atomic model that includes the locations of H or D atoms and hence provides unique information beyond that available from x-ray crystallography.

The Protein Data Bank archive is an international repository that contains more than 160 000 protein crystal structures. The majority have been determined by x-ray crystallography, which means they do not provide detailed information about the location of H atoms. Only about 150 proteins have been studied using NMX, and most of those studies occurred relatively recently as a result of developments at neutron facilities worldwide.

The lack of NMX-derived information in the data bank is a severe limitation given that H atoms are central to a protein's structure, function, and interactions. Neutrons directly visualize the locations and orientations of water molecules around biological macromolecules, the protonation of amino acids, and the details of hydrogen bond arrangements (see figure 4b).[11] In neutron density maps of crystal structures H atoms appear as negative density and D atoms appear as positive, so that deuterated proteins in which H is replaced with D greatly improve visibility. Whereas x rays allow one to see structure, neutrons allow insights related to a protein's chemistry, such as protonation shifts and charge transfer processes (see PHYSICS TODAY, November 2003, page 17).

Peter Moody at the University of Leicester and collaborators have recently done research involving the iron-containing proteins cytochrome *c* peroxidase (C*c*P) and ascorbate peroxidase (APX). That work provides an excellent example of how neutron protein crystallography has revealed novel information about the roles of protons and H_2O in enzyme catalysis. Those model systems have helped to unravel the mysteries of the degradation of hydrogen peroxide and other small molecules. C*c*P and APX carry an Fe-containing organic cofactor, known as a heme, buried in their centers. The structural arrangement of the heme is similar to that of hemoglobin, the oxygen-binding protein in blood. To carry out catalytic reactions, the heme in the center of a peroxidase needs to interact with O or an O-containing molecule such as H_2O_2.

Researchers are intrigued by the nature of Fe–O complexes—both what they look like and where the protons are located in the proteins that help stabilize them. Structures of important Fe–O complexes in C*c*P have typically been obtained by x-ray crystallography. However, the x rays release electrons that alter the chemical states of metals, including the Fe in heme enzymes. Figure 4c shows the heme portion of APX in an important Fe–O complex as determined using NMX. The image shows where the Fe–O is located and the orientation of an O–D chemical intermediate that is critical for enzymatic function. The use of NMX therefore provides new information about the structure and interactions of the heme cofactor and additional insights about how oxidative enzymes like C*c*P and APX function in a cell.

Improved instrumentation for neutron-based protein crystallography has considerably increased the speed of data acquisition and volume of data that can be obtained using NMX. Cylindrical neutron-sensitive image-plate (NIP) detectors and detector arrays maximize the solid angle of data capture and the efficiency of data acquisition. At the ILL, Matthew Blakeley and colleagues have driven successive upgrades of the LADI-III diffractometer that have delivered improvements in effective neutron flux. It now collects better-quality data with higher throughput. NMX beamlines worldwide use cryocooling to freeze short-lived chemical intermediates in protein crystals.[11] Data from those crystals enable researchers to visualize the atomic structures of unstable protein intermediates, including the locations of H atoms.

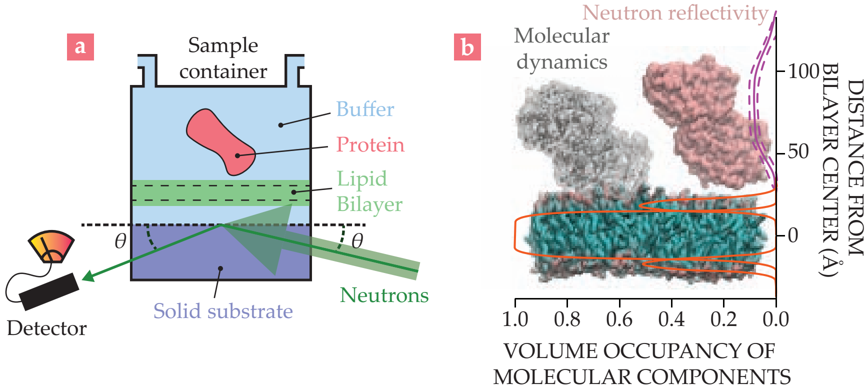


FIGURE 3. NEUTRON REFLECTOMETRY (NR) is performed on thin films of biological materials prepared on solid, flat substrates. **(a)** A typical NR sample consists of a lipid bilayer with associated biomolecules such as proteins in a fully hydrated environment, typically a sample container that allows *in situ* buffer exchange. The fraction of the beam scattered as a specular reflection—with equal incident and reflected angles *θ*—is measured over a range of angles and transformed into a reflectivity curve. **(b)** The orientation of tubulin (pink), a peripheral membrane protein, on a biomimetic mitochondrial membrane surface (blue) was derived from NR data. The results validate molecular dynamics simulations (gray) and reveal mechanistic details of both the surface binding mechanism and the motions of the surface-bound tubulin molecule.[9] (Courtesy of David Hoogerheide.)

Protein dynamics

Protein dynamics such as hinge-bending movements, large-amplitude collective motions, and sampling of different structural arrangements are essential for biological function.[12] Neutron scattering techniques are highly effective for studying motion in biological systems; they are sensitive to length scales ranging from 1 Å to 100 Å and to time scales in the femtosecond to microsecond range.

Experiments typically use solutions of proteins in various solvents that have different levels of deuteration. Protein dynamics can also be studied in living cells. Depending on the experimental configuration, data can be extracted through either incoherent or coherent scattering from samples. Techniques such as quasielastic neutron scattering (QENS) probe H dynamics in proteins by measuring incoherent scattering. Coherent scattering techniques, such as neutron spin echo (NSE) and backscattering spectrometry, detect pair correlations and can be used to study collective dynamics in proteins.[12] For example, NSE has been used to study how two large domains of the enzyme phosphoglycerate kinase move in solution to form a structure that enables it to perform catalysis.[13]

Small-angle neutron scattering supplies some information on the assembly of proteins and complexes, and information from neutron spectroscopy can add details about their movements. In the context of other biochemical and biophysical experiments, those data can elucidate macromolecular dynamics and their relationships to biological function in many processes, including enzyme catalysis, protein folding, motions in folded and intrinsically disordered proteins, and fluctuations in the ribosome structure.[13,14]

Although the number of instruments available for the study of macromolecular dynamics using neutrons is relatively small, all the major facilities have significant capability in the area and are continually upgrading and incorporating emerging technical developments. Upgrades to the ILL's backscattering and time-of-flight spectrometers have increased their incident flux and energy resolution, thus improving data quality and, crucially, reducing the amount of sample material required. J-PARC has been gradually building its user program and expanding its capability to study macromolecules since 2008; it now has six instruments for inelastic or quasielastic scattering at its Materials and Life Science Experimental Facility.

At FRM-II, inelastic scattering experiments will benefit from planned upgrades, including neutron guides with supermirror coatings and a new time-of-flight spectrometer, increasing the neutron flux to samples. The neutron spin-echo spectrometer at NIST has been upgraded with a new neutron polarization device, which has yielded a significant reduction in data collection time and required sample size while enhancing the range of accessible time scales. In 2023, with collaboration from the Center for Neutron Science at the University of Delaware and funds from NSF, a new spin-echo spectrometer will increase NIST's data collection rate by an additional order of magnitude.

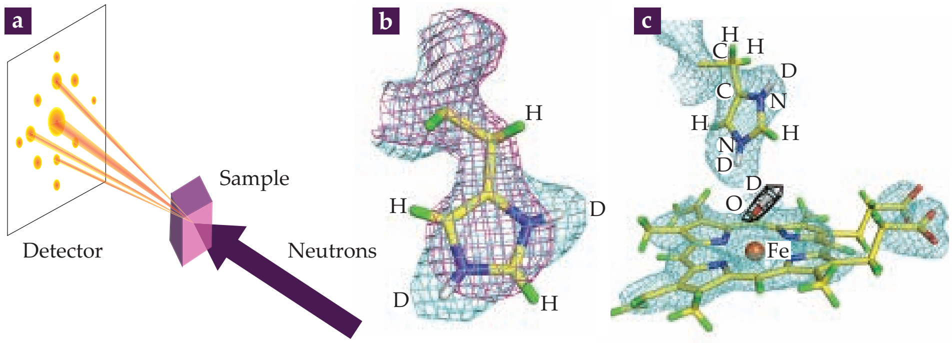


FIGURE 4. (a) IN NEUTRON MACROMOLECULAR CRYSTALLOGRAPHY (NMX), diffraction data, which are observed as reflections, are collected from a deuterated water–soaked protein crystal sample irradiated by a neutron beam. Reflections are transformed into a nuclear scattering density using phase information, which is often obtained from the x-ray crystal structure of the same protein. **(b)** The purple mesh shows the electron density of a histidine residue; the cyan mesh shows the corresponding density derived from neutron data. The atomic structure is shown as a stick diagram with hydrogen atoms in green and deuterium in white. The nonoverlapping cyan mesh indicates the presence of D atoms that can only be visualized using neutrons. **(c)** A portion of the crystal structure of ascorbate peroxidase (APX) was determined by combining crystallographic data obtained using x rays and neutrons. The stick diagram shows the atomic structure; cyan mesh shows the density from the neutron map surrounding the iron-containing compound (orange sphere) in APX. The black mesh depicts the nuclear density of an oxygen–deuterium chemical intermediate (red and white stick). Studying the interactions of the O–D intermediate with an Fe-containing compound helps uncover how oxidative enzymes function. (Panel a courtesy of David Hoogerheide; panels b and c adapted by Peter Moody at the University of Leicester, UK, from ref. 11, H. Kwon et al.)

Emerging capability

Neutron facilities worldwide are working to increase the range and diversity of possible structure-based studies of biological systems. New diffractometers are being developed to improve NMX and to reduce demand on existing instruments. The ILL is constructing a new instrument called DALI to improve capacity and to allow researchers to study crystals with larger unit cells. New, fast detectors on the Macromolecular Neutron Diffractometer at ORNL (shown in figure 5a), a biological crystal diffractometer at J-PARC, and the possible use of a gas electron multiplier detector at the ESS will increase those facilities' capabilities.

As throughput increases, so will the need for data portability and standardization across sources and beamlines. The Diffraction Integration for Advanced Light Sources project has developed open-source software to address that need by producing a framework for the collection and analysis of data.

One example of new instrumentation for NR is the Chromatic Analysis Neutron Diffractometer or Reflectometer, which is being commissioned at NIST. Shown in figure 5b, CANDOR is a white-beam reflectometer featuring a novel scintillator-based, position-sensitive and energy-selective detector. It combines energy discrimination—the primary advantage of pulsed-source reflectometry—with the uninterrupted flow of neutrons from a continuous source. By utilizing a large range of neutron wavelengths, CANDOR is expected to increase the neutron flux at the sample position by more than an order of magnitude compared with current instruments.

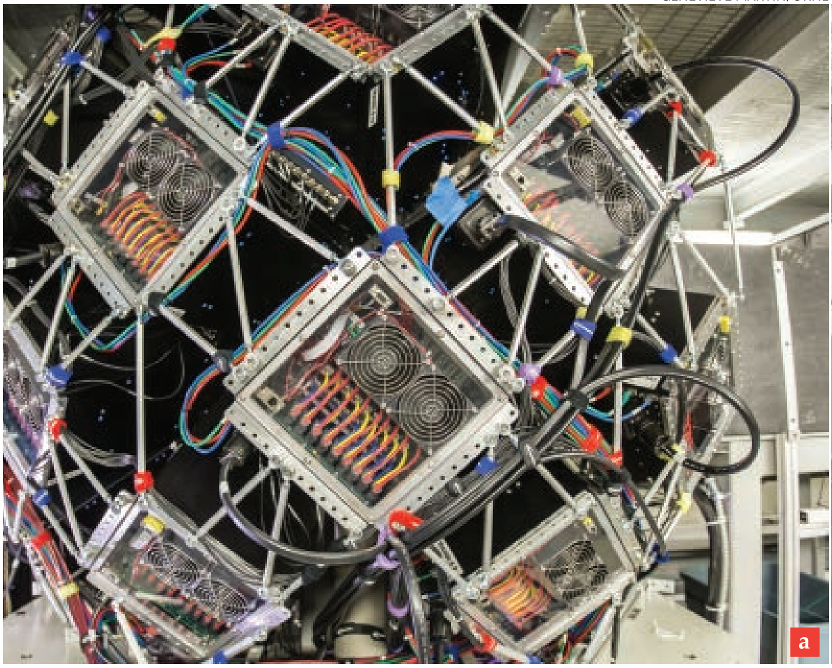
Neutrons have considerable potential for imaging on length scales of multicellular organizations and tissues. They are particularly well-suited to that job because their penetration depth into biological materials can be tuned by changing the irradiation energy. For example, neutron tomography techniques, including emerging methods such as small-angle tensor tomography,[15] offer the possibility of imaging 3D structures in tissues and whole organisms with micrometer resolution. The ESS has identified neutron tomography as a technique that could benefit from its higher brilliance source, which allows for more rapid data collection and improved resolution.

Other recent innovations include axisymmetric mirrors to reflect and focus neutrons. Wolter optics, originally developed for telescopes, have already been implemented in a compact SANS instrument at the Spallation Neutron Source at ORNL.[16] Such setups are opening possibilities for neutron imaging of live biological samples.

Strategies and collaboration

In July 2018 the American Physical Society's Panel on Public Affairs published a report, *Neutrons for the Nation*. The 32-page report highlights current and future issues around the need to maintain national facilities supporting R&D that involves neu-

GENEVIEVE MARTIN/ORNL DAN NEUMANN/NIST



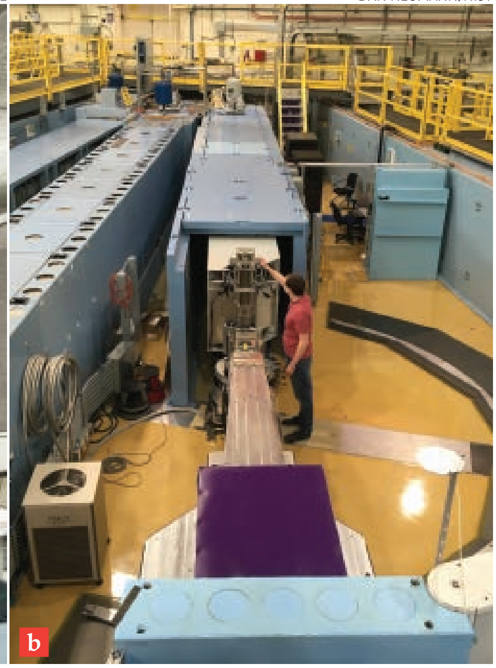


FIGURE 5. (a) THE LARGE METAL FRAMEWORK OF THE MACROMOLECULAR NEUTRON DIFFRACTOMETER (MaNDi) at Oak Ridge National Laboratory supports numerous scintillation detectors used to collect diffraction data from protein crystals. **(b)** The new Chromatic Analysis Neutron Diffractometer or Reflectometer (CANDOR) at the NIST Center for Neutron Research in Gaithersburg, Maryland, discriminates between neutron energies using Bragg diffraction with highly oriented pyrolytic graphite analyzer crystals. MaNDi and CANDOR represent a new generation of instruments for collecting neutron data with increased sensitivity and higher throughput.

trons. Such activities are "essential components of R&D in numerous areas of science and engineering," the report says. "Neutron sources play a key role in overall U.S. innovation capacity."

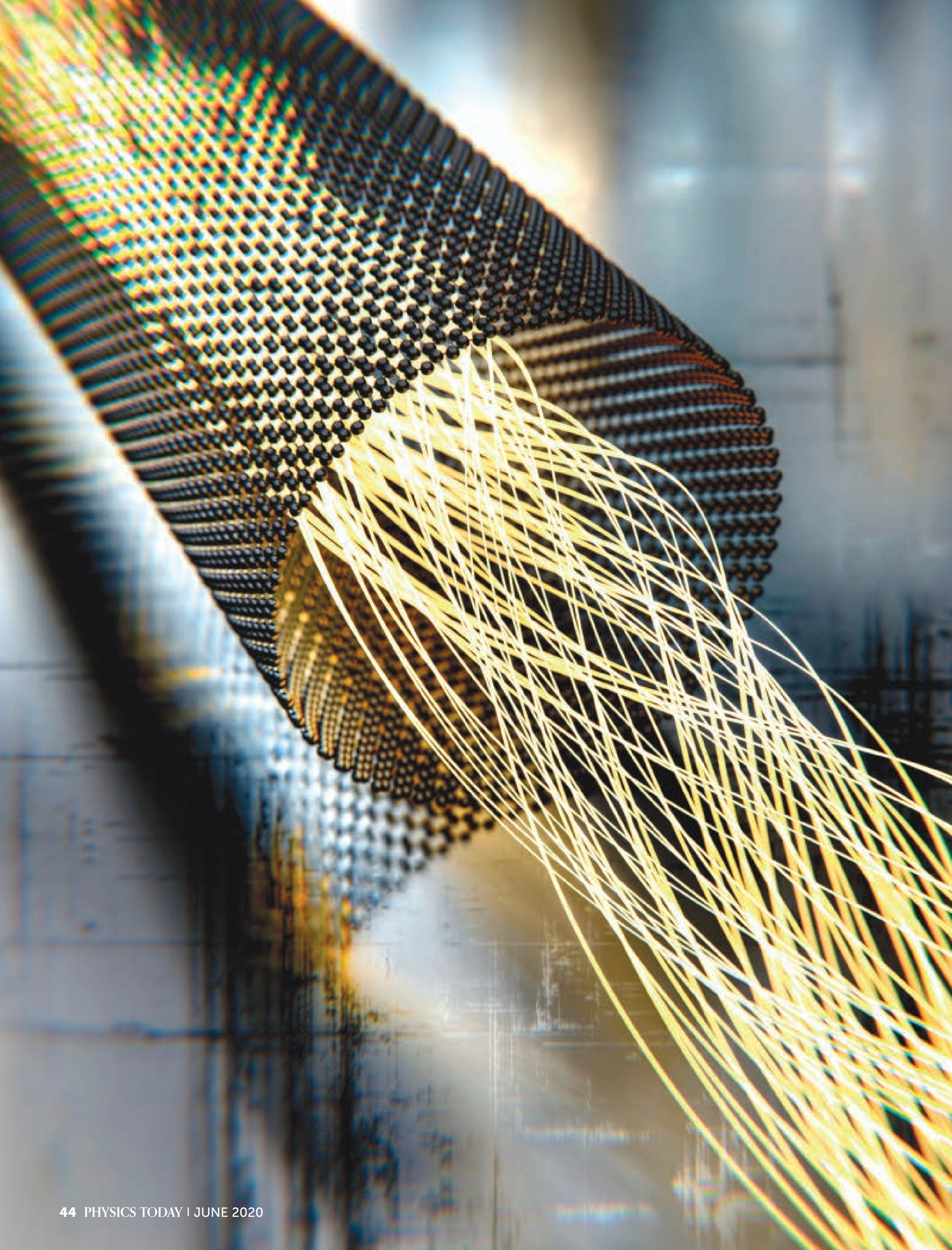
The report also points to a decline in neutron R&D efforts in the US. It acknowledges an urgent need to invest in maintaining diverse and complementary capabilities, including spallation facilities, research reactors, and high-performance instrumentation, as part of reestablishing the US as a world leader in neutron research. In Europe, the League of Advanced European Neutron Sources was created to coordinate efforts to increase the impact of neutron science. The key aim of the league's charter, signed in September 2018, is to "facilitate any form of discussion and decision-making process that has the potential to strengthen European neutron science via enhanced collaboration among the facilities."

Well-defined strategies and cooperation between facilities at national and international levels will help maximize the potential for using neutrons to study biological problems. An emerging trend among large-scale facilities is to give users access to multiple techniques, such as state-of-the-art neutron, x-ray, cryoelectron microscopy, and NMR platforms, at or near the same site, as now routinely happens at the Grenoble Partnership for Structural Biology. Biological neutron scattering is at its most powerful when combined with other techniques that provide unique insights into the structure, assembly, and function of macromolecules. Such collaboration presents a great opportunity for physicists to engage with challenging problems in biological research to impact both fundamental knowledge and human health.

REFERENCES

1. V. T. Forsyth, P. Moody, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **74**, 1126 (2018); R. Ashkar et al., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **74**, 1129 (2018); G. Zaccai, *Eur. Biophys. J.* **41**, 781 (2012).
2. J. H. Lakey, *J. R. Soc. Interface* **6**, S567 (2009); R. Schneider et al., *J. Mol. Biol.* **41**, 231 (1969); H. B. Stuhrmann, *J. Appl. Crystallogr.* **7**, 173 (1974); B. P. Schoenborn, *Nature* **224**, 143 (1969); B. Jacrot, *Rep. Prog. Phys.* **39**, 911 (1976).
3. M. Haertlein et al., in *Isotope Labeling of Biomolecules—Applications*, Z. Kelman, ed., Academic Press/Elsevier (2016), p. 113.
4. A. Lapinaite et al., *Nature* **502**, 519 (2013).
5. S. Maric et al., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **70**, 317 (2014); I. Josts et al., *Structure* **26**, 1072 (2018); J. Nitsche et al., *Commun. Biol.* **1**, 206 (2018).
6. E. Mahieu, F. Gabel, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **74**, 715 (2018); J. Trewhella, in *Integrative Structural Biology with Hybrid Methods*, H. Nakamura et al., eds., Springer (2018), p. 77; A. Jordan et al., *J. Appl. Crystallogr.* **49**, 2015 (2016); E. Valentini et al., *Nucleic Acids Res.* **43**, D357 (2015).
7. S. Grudinin, M. Garkavenko, A. Kazennov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **73**, 449 (2017).
8. C. F. Majkrzak et al., in *Neutron Scattering in Biology: Techniques and Applications*, J. Fitter, T. Gutberlet, J. Katsaras, eds., Springer (2006), p. 225.
9. P. Shekhar et al., *J. Appl. Phys.* **110**, 102216 (2011); D. P. Hoogerheide et al., *Proc. Natl. Acad. Sci. USA* **114**, E3622 (2017).
10. S. Waldie et al., *Langmuir* **34**, 472 (2018); M. Moulin et al., *Chem. Phys. Lipids* **212**, 80 (2018); A. Luchini et al., *Colloids Surf. B* **168**, 126 (2018); S. Waldie et al., *Sci. Rep.* **9**, 5118 (2019).
11. C. M. Casadei et al., *Science* **345**, 193 (2014); H. Kwon et al., *Nat. Commun.* **7**, 13445 (2016); M. G. Cuypers et al., *Sci. Rep.* **6**, 31487 (2016); A. W. Yee et al., *Nat. Commun.* **10**, 925 (2019); G. Zaccai, *Biochim. Biophys. Acta* **1864**, 129475 (2020).
12. J. C. Smith et al., *Annu. Rev. Biophys.* **47**, 335 (2018).
13. R. Inoue et al., *Biophys. J.* **99**, P2309 (2010).
14. M. R. Duff Jr et al., *Biochemistry* **57**, 4263 (2018); F. Ameseder et al., *Phys. Chem. Chem. Phys.* **20**, 5128 (2018); G. Schirò et al., *Nat. Commun.* **6**, 6490 (2015); G. Zaccai et al., *Sci. Rep.* **6**, 37138 (2016).
15. M. Liebi et al., *Nature* **527**, 349 (2015).
16. D. Liu et al., *Nat. Commun.* **4**, 2556 (2013).

PT



ISTOCK.COM/NIPLOT

Ramón Aguado is a senior researcher at the Spanish National Research Council (CSIC) in Madrid. **Leo Kouwenhoven** is a researcher at the Microsoft Quantum Lab Delft and a professor of applied physics at Delft University of Technology in the Netherlands.

Majorana qubits for topological quantum computing

Ramón Aguado and Leo P. Kouwenhoven

Researchers are trying to store robust quantum information in Majorana particles and are generating quantum gates by exploiting the bizarre non-abelian statistics of Majorana zero modes bound to topological defects.

Soon after Enrico Fermi became a professor of physics at Italy's University of Rome in 1927, Ettore Majorana joined his research group. Majorana's colleagues described him as humble because he considered some of his work unexceptional. For example, Majorana correctly predicted in 1932 the existence of the neutron, which he dubbed a neutral proton, based on an atomic-structure experiment by Irène Joliot-Curie and Frédéric Joliot-Curie. Despite Fermi's urging, Majorana didn't write a paper. Later that year James Chadwick experimentally confirmed the neutron's existence and was awarded the 1935 Nobel Prize in Physics for the discovery.

Nevertheless, Fermi thought highly of Majorana, as is captured in the following quote: "There are various categories of scientists, people of a secondary or tertiary standing, who do their best but do not go very far. There are also those of high standing, who come to discoveries of great importance, fundamental for the development of science. But then there are geniuses like Galileo and Newton. Well, Ettore was one of them." Majorana only wrote nine papers, and the last one, about the now-eponymous fermions, was published in 1937 at Fermi's insistence. A few months later, Majorana took a night boat to Palermo and was never seen again.[1]

In that final article, Majorana presented an alternative representation of the relativistic Dirac equation in terms of real wavefunctions. The representation has profound consequences because a real wavefunction describes particles that are their own antiparticles, unlike electrons and positrons. Since particles and antiparticles have opposite charges, fermions in his new representation must have zero charge. Majorana postulated that the neutrino could be one of those exotic fermions.

Although physicists have observed neutrinos for more than 60 years, whether Majorana's hypothesis is true remains unclear. For example, the discovery of neutrino oscillations, which earned Takaaki Kajita and Arthur McDonald the 2015 Nobel Prize in Physics, demonstrates that neutrinos have mass. But the standard model requires that neutrinos be massless, so various possibilities have been hypothesized to explain the discrepancy. One answer could come from massive neutrinos that do not interact through the weak nuclear force. Such sterile neutrinos could be the particles that Majorana predicted. Whereas conclusive evidence for the existence of Majorana neutrinos remains elusive, researchers are now using Majorana's idea for other applications, including exotic excitations in superconductors.

Majorana quasiparticles in superconductors

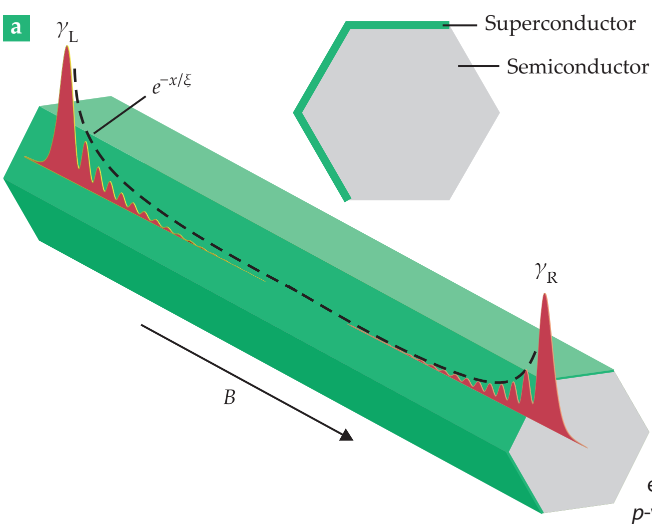
From the condensed-matter viewpoint, Majoranas are not elementary particles but rather emergent quasiparticles. Interestingly, the equation that describes quasiparticle excitations in superconductors has the same mathematical structure as the Majorana equation. The reason for the similarity arises from the underlying particle–hole symmetry in superconductors: Unlike quasiparticles in a metal, which have a well-defined charge, quasiparticles in a superconductor comprise coherent superpositions of electrons and holes. For the special zero-energy eigenmode, the electron and the hole, which each contribute half probability, form a quasiparticle. The operators describing the zero-energy particle–hole superpositions are invariant under charge conjugation, and zero-energy modes are therefore condensed-matter Majorana particles.

Particle–hole symmetry dictates that excitations in superconductors should occur in pairs at energies $\pm E$. Therefore, zero-energy excitations are seemingly unreachable because they cannot emerge by any smooth deformation of the Hamiltonian, which would require that one of the solutions disappear. Rather, the only way to generate zero-energy excitations in superconductors is through a topological transition, a process that separates the phase of Majorana zero modes from the phase without them by closing and then reopening the superconducting gap (see the article by Nick Read, PHYSICS TODAY, July 2012, page 38).

Majorana zero modes are located at topological defects, such as vortices, boundaries, and domain walls in topological superconductors. Remarkably, Majorana zero modes bound to defects do not obey fermion statistics. Unlike the original particles predicted by Majorana, the zero modes possess non-abelian exchange statistics, also known as non-abelian braiding, which makes them promising for applications in topological quantum computing, as detailed in box 1. Quasiparticles with non-abelian exchange statistics were first predicted in 1991 to occur in the filling factor $\nu = 5/2$ of the fractional quantum Hall state. In 2000, researchers demonstrated that similar physics occur in superconductors with intrinsic *p*-wave pairing, an exotic form of superconductivity in which Cooper pairs bind through rare triplet-like pairing instead of the more standard singlet-like pairing in *s*-wave superconductors.[2] Conventional *s*-wave pairing can be converted to *p*-wave pairing by combining the superconducting proximity effect in materials with strong spin–orbit interactions and an external magnetic field that breaks time-reversal symmetry.

The nanowire proposal

In 2010 two research groups made an elegant theoretical proposal, shown schematically in figure 1. If a semiconducting nanowire with strong spin–orbit coupling, such as indium arsenide or indium antimonide, is coupled to a standard *s*-wave superconductor, Majorana zero modes will emerge at both ends of the nanowire, provided that a magnetic field is applied parallel to it.[3] The proposal realistically implements the paradigmatic one-dimensional model for *p*-wave superconductivity



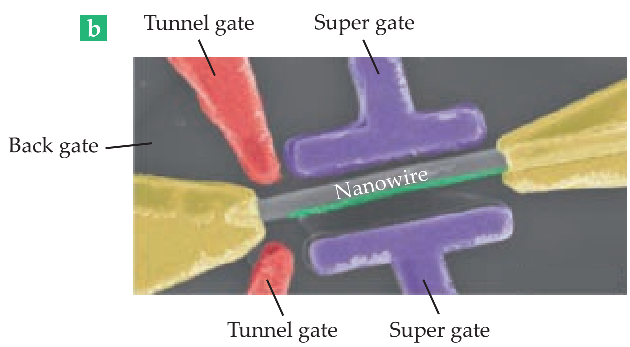


FIGURE 1. (a) THE NANOWIRE PROPOSAL[3] takes a nanowire of a semiconductor, such as indium arsenide or indium antimonide, that has strong spin–orbit coupling and places it in contact with an *s*-wave superconductor, such as aluminum, in the presence of an external magnetic field *B*. As in the original model for one-dimensional *p*-wave superconductors,[4] the nanowire device experiences a topological nontrivial phase with exponentially decaying Majorana bound states, denoted γ_L, at both ends of the nanowire. **(b)** An actual device from Delft University of Technology includes various metallic gates for tuning it to the topological phase by adjusting the nanowire's chemical potential. (Panel a adapted from ref. 3, R. M. Lutchyn, J. D. Sau, S. Das Sarma; panel b adapted from H. Zhang et al., *Nature* **556** 74, 2018.)

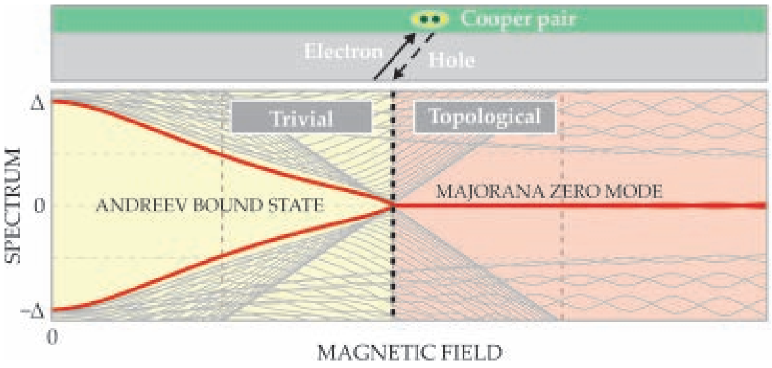


FIGURE 2. ANDREEV REFLECTIONS of electrons and holes to form Cooper pairs at the semiconducting–superconducting interface induce superconductivity in a nanowire. As a result, Majorana zero modes (flat red line) emerge in the energy spectrum as the external magnetic field increases. The Majoranas appear beyond some critical value of the external field (black dotted line) where the superconducting gap closes and reopens again, which signals a topological phase transition. Theory predicts that the emergent Majorana zero modes can be detected as a zero-bias anomaly in electrical conductance dI/dV. (Image by R. Aguado and L. P. Kouwenhoven.)

that was discussed in 2001 for the first time by Alexei Kitaev.[4]

The Majorana zero modes are localized at opposite ends of the wire and decay with position x as $e^{-x/\xi}$, where ξ is the localization length. But together they form a highly delocalized fermion, which can be seen mathematically as a fermion operator that decomposes to two real, self-adjoint operators. The nonlocal fermion defines two parity states—the empty state and the full fermion one—that are degenerate at zero energy except for exponentially small corrections $e^{-L/\xi}$, where L is the length of the wire. Those two states can be used to define a qubit. Because the states are stored nonlocally, the qubit is resilient against local perturbations from the environment.

To induce a closing and reopening of an energy gap in the nanowire platform, researchers exploit the competition among three effects. The first, the s-wave superconducting proximity effect, pairs electrons of opposite spin and opens a superconducting gap Δ at the Fermi level. In the second effect, an external magnetic field B generates a Zeeman energy $E_Z = g\mu_B B/2$—with g the nanowire's Landé factor and μ_B the Bohr magneton—which tends to break Cooper pairs by aligning their electron spins and closing the gap. The third effect, spin–orbit coupling, negates the external magnetic field by preventing the spins from reaching full alignment.

The competition between the second and third effects creates regions in parameter space where the gap closes and reopens again. At low electron densities, the transition occurs when the Zeeman energy is of the same magnitude as the induced superconducting gap, and it can be reached either by increasing the magnetic field, as shown in figure 2, or by tuning the wire's chemical potential. Apart from choosing semiconductors with a large spin–orbit coupling and good proximity effect with conventional superconductors, researchers need large g factors to induce a large Zeeman effect with moderate magnetic fields below the critical field of the superconductor. Materials such as the heavy-element semiconductors InAs and InSb have proven to be excellent choices.

Topological superconductivity can also be engineered using similar ideas in alternative platforms. Some examples include chains of magnetic impurities above superconductors; proximitized 2D materials; and vortices in proximitized topological insulators such as quantum spin-Hall insulators, quantum anomalous-Hall insulators, and iron-based topological surface states.

Measuring Majoranas

At energies below the superconducting gap, an electron incident on a superconductor (S) from a normal conductor (N) can be reflected either as an electron or as a hole. Whereas the electron process is a standard, normal reflection, the hole process, known as Andreev reflection, is subtler because electrons are reflected as holes in the normal side while creating a Cooper pair in the superconducting side. In a standard NS junction, such Andreev

BOX 1. NON-ABELIAN BRAIDING

Quantum mechanics dictates that particles obey either Fermi–Dirac or Bose–Einstein statistics in three dimensions, which means that the wavefunction Ψ of a system of indistinguishable particles is necessarily bosonic or fermionic upon particle exchange. From that point of view, fermions and bosons are not exotic because exchanging them leaves the ground state invariant, up to a sign: $\Psi \to \pm\Psi$.

Two dimensions are richer. Now, the possibilities go beyond the fermionic or bosonic cases. A system can exhibit anyon statistics in which the wavefunction picks up an arbitrary phase under an exchange: $\Psi \to e^{i\theta}\Psi$. Such behavior generalizes the boson and fermion cases, where the phases can only be $\theta = 0$ or $\theta = \pi$. Because the phase factors are ordinary commuting numbers, the order of successive exchanges doesn't matter, and the anyon statistics are called abelian.

The weirdness starts in systems with a degenerate many-body ground state containing several quasiparticles. When quasiparticles are exchanged, the system goes from one ground state, Ψ_a, to another, $M_{ab}\Psi_b$. Because the unitary transformations M_{ab} that operate in the subspace of degenerate ground states are generally noncommuting, the anyonic statistics take a non-abelian form. The final state of the system, therefore, depends on the order of the exchange operations, similar to braiding cords in a necklace.

Using Majorana zero modes to store and manipulate quantum information is one case where non-abelian braiding statistics form the basis of topological quantum computation (see the article by Sankar Das Sarma, Michael Freedman, and Chetan Nayak, PHYSICS TODAY, July 2006, page 32). Quantum computation in such a system also benefits from protection against environmental decoherence because of the nonlocal character of Majorana-based qubits.

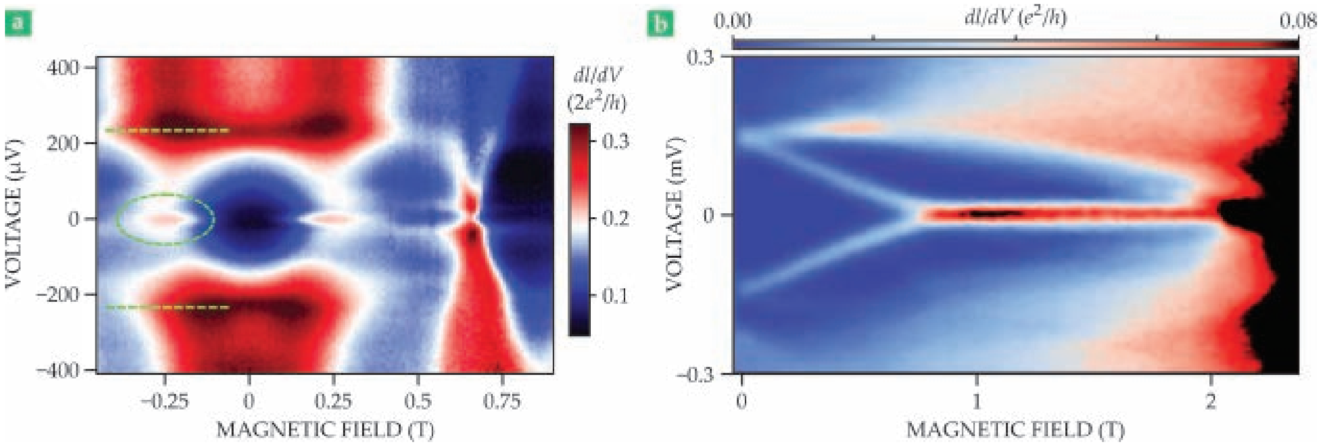


FIGURE 3. (a) A CONTOUR PLOT of dI/dV versus voltage V and external magnetic field B along the axis of an indium antimonide nanowire in contact with niobium titanate nitrate shows that for fields between 100 mT and 400 mT, a clear zero-bias anomaly (ZBA, green dotted oval) emerges in the middle of the superconducting gap, denoted by the dashed green lines. **(b)** An indium arsenide nanowire coupled to aluminum shows a robust ZBA. (Panel a adapted from ref. 5; panel b adapted from ref. 8.)

processes are rare in the tunneling limit, and the conductance is small. But in a topological NS junction containing Majorana bound states, an incident electron is always reflected as a hole with unitary probability.

As a result of that resonant Andreev process, the electrical conductance G at zero voltage is expected to be perfectly quantized: $G=2e^2/h$, where e is the electron charge and h, Planck's constant. The Andreev process underscores the particle–antiparticle duality of Majorana bound states: Because the electron and hole contribute equally to form a Majorana quasiparticle, the tunneling rates for electrons and holes should be equal. Therefore, researchers can use tunneling spectroscopy to directly detect a Majorana bound state as a zero-bias anomaly (ZBA). The differential conductance dI/dV, with I the current across the junction, is a function of the applied bias voltage V, and the ZBA should emerge as an increasing magnetic field induces a topological transition in the nanowire.

In 2012, researchers showed that the nanowire proposal could indeed be realized.[5] A typical measurement from that experiment is illustrated in figure 3a, which shows conductance versus applied bias voltage and magnetic field. For intermediate values of the magnetic field, a clear ZBA emerges in the middle of the superconducting gap and is consistent with the existence of zero-energy Majorana bound states in the nanowire. Subsequent experiments showed similar results.[6]

The debate

Members of the research community greeted the nanowire experiments with excitement (see PHYSICS TODAY, June 2012, page 14), and they also challenged the Majorana interpretation. Many features of the experiments, notably the absence of a closing and reopening of the gap and a conductance well below the quantized $G=2e^2/h$ limit, disagreed with model predictions. Importantly, a similar ZBA unrelated to Majoranas may also appear because of various physical mechanisms, such as the Kondo effect and disorder. Those are related to the sizeable subgap conductance that arises from an imperfect superconducting proximity effect.

Fortunately, many of the false-positive scenarios can now be ruled out because of advances in materials and fabrication. Some examples include the epitaxial growth of crystalline superconductor shells directly on the surface of the nanowires and the careful engineering of high-quality semiconductor–superconductor interfaces.[7] That progress has generated improved devices with much better induced superconductivity, including negligible subgap conductance. The new generation of devices has produced cleaner data with robust ZBAs,[8] as shown in figure 3b, and values close to the expected $G=2e^2/h$ ideal limit.[9]

Can we now claim that Majoranas have been observed? We cannot, because topological protection has not yet been demonstrated. Variations in the electrostatic potential, including from disorder and inhomogeneous gating, can produce regions where Andreev levels—the superconducting counterparts of particle-in-a-box confined states in quantum mechanics—appear at zero energy without a concomitant topological transition. Theory predicts that those zero modes are ubiquitous.[10] Physically, they correspond to single-fermionic Andreev levels that can be decomposed into two Majoranas. Because they partially overlap in space, those Majoranas lack the full topological protection offered by spatial separation. The tunneling coupling to the normal conductor can be distinct, which results in robust nontopological ZBAs. Even without a topological phase in the nanowire, those Majoranas still obey non-abelian statistics.

Now the challenge is to demonstrate that Majorana zero modes can be generated with an exponentially small overlap such that deviations from perfect ground-state degeneracy are exponentially small. In that regime, Majoranas become topologically protected and can be used to define parity qubits. Recent efforts to extract the degree of Majorana nonlocality have started to appear in the literature.[11] Researchers have also made experimental advances with the superconductor–semiconductor interface, among them a thin aluminum layer epitaxially covering a high-mobility InAs 2D electron gas and an Al shell wrapping an InAs nanowire core. Both schemes represent a paradigm shift because they allow topological superconductivity to be tuned by controlling the phase of the superconducting order parameter rather than by the Zeeman effect. Researchers can also

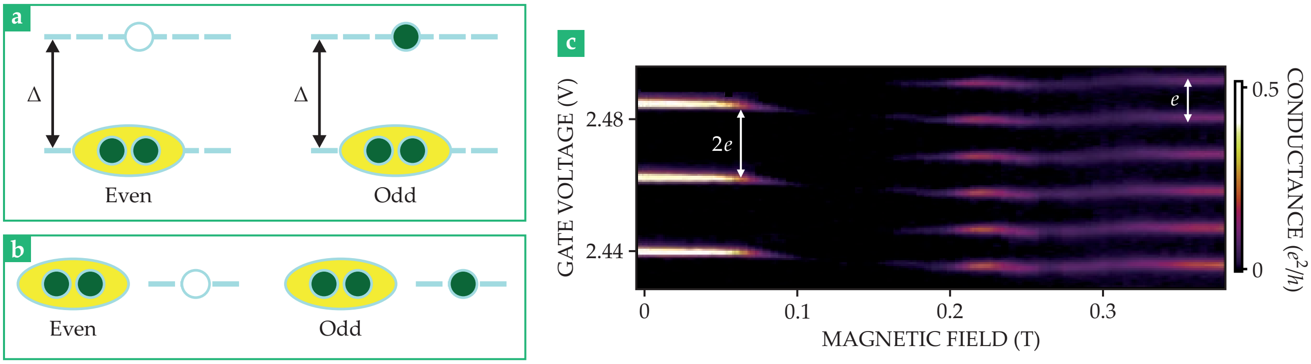


FIGURE 4. MAJORANA ISLANDS are based on finite-sized nanowire segments in a floating geometry and placed in contact with superconductors. **(a)** The energy cost for adding an extra electron in a standard superconductor is given by the superconducting gap Δ. When the gap is larger than the Coulomb energy to charge the island, the system can only accommodate electrons in pairs. **(b)** In a topological superconductor, single electrons can be accommodated at no energy cost by filling the zero-energy fermionic state formed by two nonlocal Majoranas. **(c)** Observed linear conductance data are graphed as a function of gate voltage for increasing magnetic field. The series of 2*e*-periodic Coulomb blockade peaks at low magnetic fields become 1*e*-periodic for larger magnetic fields. (Adapted from ref. 14.)

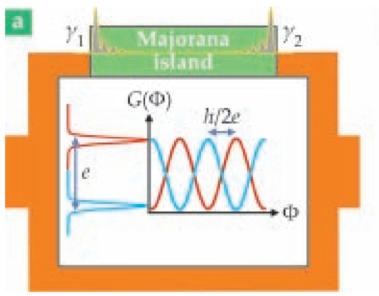
use those schemes to control Josephson junctions fabricated with a 2D electron-gas hybrid material[12] and even full phase windings, akin to vortices, in the full-shell nanowire geometry.[13]

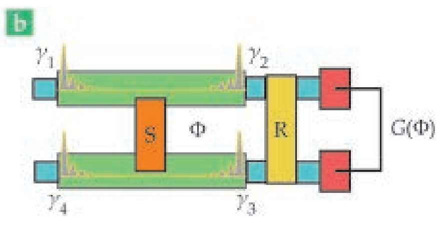
Majorana islands

Another nanowire option called the floating geometry electrically isolates a nanowire with small capacitors. That geometry produces Majorana islands, which show deviations from ground-state degeneracy. The scheme exploits even–odd effects in small superconductors. Recall that the classical energy to charge a capacitor is inversely proportional to its capacitance; for a sufficiently small island, the charging energy for adding a single extra electron can be significant. Transport through such an island is blocked at low voltages and temperatures, a phenomenon called Coulomb blockade. Current flow—so-called Coulomb blockade peaks—is possible only at special degeneracy points. They occur periodically at gate voltages for which the energy of having N or $N+1$ electrons on the island is equal. In the presence of superconductivity, Coulomb blockade still applies, though the energy of N electrons depends also on fermionic parity. If N is even, all quasiparticles couple as Cooper pairs in the ground state. Adding an extra electron costs both electrostatic charging energy and a finite energy that corresponds to the lowest quasiparticle excitation. For a standard superconductor, the energy of the odd-N configuration is the superconducting gap, as shown in figure 4a.

BOX 2. MAJORANA-BASED QUBITS

Researchers have explored various schemes that use two-path electron interferometry for parity readout. One path involves a Majorana island; the other serves as a reference. Part a of the figure shows an interferometer with a Majorana island (green) in the Coulomb blockade regime. In that setup, the amplitudes of the Coulomb blockade conductance peaks display Aharonov–Bohm oscillations as a function of the external magnetic flux Φ piercing the interferometer; the oscillations reflect coherent single-electron transport—electron teleportation—across the island. The plot of the interferometer conductance G (Φ) against Φ shows the oscillations of two successive Coulomb peaks. The π phase shift denotes a change in the fermion parity of the island.[16]





Proposals for performing parametric braiding, rather than spatial braiding, rely on the interferometer. That reliance avoids the need to move Majoranas around each other in complicated geometries, such as T-junctions.[17] In a recent paper on parametric braiding, researchers have proposed the possibility of performing measurement-based topological quantum computation with quantum gates based on interferometers.[18] In the schematic shown in part b of the figure, two topological superconductor nanowires (green) are shunted by a superconducting bridge S (orange) to form a floating island hosting four Majorana modes. A reference arm R (yellow) is shorter than the coherence length and closes the interference loop. The measured flux-dependent conductance depends on the fermion parity of Majoranas 2 and 3. Such box qubits allow for measurement-only protocols, including qubit readout for the three Pauli operators in the Majorana basis and full one-qubit control using tunable couplings between Majorana states and quantum dots. (Adapted from D. Aasen et al., *Phys. Rev. X* **6**, 031016 (2016).)

If the gap is larger than the charging energy, however, electrons enter the island only as Cooper pairs. As the magnetic field increases, the quasiparticle gap reduces until the superconductor becomes topological. At that point, it can accommodate one extra electron in the nonlocal Majorana zero-mode state with zero energy, regardless of whether it is empty or full, as figure 4b illustrates. Experiments with short proximitized nanowires show a change in periodicity, evident in figure 4c, that is consistent with Majorana theory: The flow of Cooper pairs transitions to single electrons as the magnetic field increases.[14]

Toward Majorana-based qubits

The Coulomb blockade peaks of Majorana islands have a maximum conductance of $G = e^2/h$, half that of noninteracting wires, because two-charge transfers are strongly suppressed. Only single-electron tunneling is possible at charge-degeneracy points. The transfer occurs through the fermion state formed by two distant Majoranas.[15] Researchers can use that nonlocal resonant process for parity readout. And as detailed in box 2, the process enables the non-abelian braiding of Majoranas and the building of Majorana-based qubits.

Two key steps need to be completed, however, before Majorana qubits can be used in topological quantum computing. Researchers first must establish unambiguously that in the lab they can make fully nonlocal Majoranas with the requisite topological protection, demonstrate ground-state degeneracy, and test for simple measurements, such as fusion rules. The second important step is to demonstrate parity-dependent interferometry, which is at the heart of measurement-only Majorana box qubits. And on the way to topological quantum computing, researchers can explore more exotic physics, including the topological Kondo effect. That phenomenon exploits the analogy between the two degenerate parity ground states formed by two highly nonlocal Majoranas and the spin-½ system of the standard Kondo effect. Each of those milestones would help lay the foundation for topological quantum computing. By themselves, the milestones would represent an unprecedented advancement for fundamental physics.

REFERENCES

1. E. Recami, *Int. J. Mod. Phys. D* **23**, 1444009 (2014).
2. G. Moore, N. Read, *Nucl. Phys. B* **360**, 362 (1991).
3. R. M. Lutchyn, J. D. Sau, S. Das Sarma, *Phys. Rev. Lett.* **105**, 077001 (2010); Y. Oreg, G. Refael, F. von Oppen, *Phys. Rev. Lett.* **105**, 177002 (2010).
4. A. Y. Kitaev, *Phys.-Usp.* **44**, 131 (2001).
5. V. Mourik et al., *Science* **336**, 1003 (2012).
6. R. Aguado, *Riv. Nuovo Cimento* **40**, 523 (2017).
7. S. Gazibegovic et al., *Nature* **548**, 434 (2017).
8. M. T. Deng et al., *Science* **354**, 1557 (2016).
9. F. Nichele et al., *Phys. Rev. Lett.* **119**, 136803 (2017); F. Setiawan et al., *Phys. Rev. B* **96**, 184520 (2017).
10. E. Prada et al., https://arxiv.org/abs/1911.04512.
11. M.-T. Deng et al., *Phys. Rev. B* **98**, 085125 (2018).
12. A. Fornieri et al., *Nature* **569**, 89 (2019).
13. S. Vaitiekėnas et al., *Science* **367**, eaav3392 (2020).
14. S. M. Albrecht et al., *Nature* **531**, 206 (2016).
15. L. Fu, *Phys. Rev. Lett.* **104**, 056402 (2010).
16. A. M. Whiticar et al., https://arxiv.org/abs/1902.07085.
17. J. Alicea et al., *Nat. Phys.* **7**, 412 (2011).
18. S. Vijay, L. Fu, *Phys. Rev. B* **94**, 235446 (2016); S. Plugge et al., *New J. Phys.* **19**, 012001 (2017); T. Karzig et al., *Phys. Rev. B* **95**, 235305 (2017).

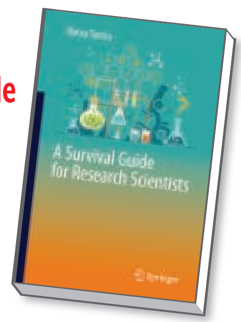
ISTOCK.COM/AKINDO

Learning from the voice of research experience

From the moment I saw its cover, I was suspicious about this book's title. Was *A Survival Guide for Research Scientists* another abominable student self-help book? My first reaction on opening the book and seeing its unusual format was even worse. Goodness, what is this—a compendium of PowerPoint presentations? A salad of disconnected paragraphs? A barrage of spider diagrams and bullet points?

I am often reminded of psychologist Daniel Kahneman's description of our cognitive dichotomy. We all experience initial reactions that are quick, harsh, and crude; we are also capable of later reactions that are slow, rational, and balanced. I know that unfortunately my initial reaction is wrong 60–70% of the time, so I suspended judgment and decided that my calmer thought processes should have the chance to weigh in on

A Survival Guide for Research Scientists
Ratna Tantra
Springer, 2019.
$109.99 (paper)



the book. After half an hour, I started to mentally apologize to the author. Ratna Tantra, an expert in microfluidics and nanobiosystems who has experience working in both industrial and government settings, has produced a remarkable piece of general advice for research scientists. Everything in this little book is useful.

That is not to say that I agree with her on every point, or on matters of writing style—I hate bullet points. But in 21 chapters, *A Survival Guide for Research Scientists* discusses with clarity, intelligence, and practical sense the most pressing troubles of a research career. I wish I had read it decades ago. The book's tips on navigating the laboratory, writing reports and proposals, conducting meetings and interviews, working in teams, and relating to difficult collaborators reflect the hard-won lessons of an experienced researcher.

Most chapters refer to the author's experiences, a choice that gives intimacy and freshness to the narrative. I particularly enjoyed the comments on frequent reasons why research proposals are rejected, including ineligibility for funding, weak science, inexperienced research team, lack of credibility, unclear proposal, unrealistic budget, and little added value to the existing science. That litany is familiar to me and most other research scientists. Also, the chapters on dealing with layoffs and self-employment are extremely timely.

The first part of the book deals with self-care. The author discusses topics including stress, anxiety, personal problems, and work–life balance. I initially doubted Tantra's decision to start with this matter. But after reading Chris Woolston's 2019 report in *Nature* (volume 575, page 403) about mental health in PhD students, I realized that genuine "epidemics of stress" loom on university campuses and that graduate students are particularly vulnerable. *Nature*'s international survey of 6300 doctoral students found that 71% of respondents were satisfied with their experience of research, but 36% sought help for anxiety or depression related to their PhD. In Advance HE's national survey of graduate students in the UK, 86% of the 50 000 respondents reported levels of anxiety much higher than those of the general population.

What causes such high levels of stress, anxiety, and depression among young research scientists? Long work hours, financial troubles, poor leadership and mentoring, bullying and harassment, discrimination, intellectual and technological overload, and poor job prospects all contribute to declining mental health. In my opinion, scientists' jobs have also become vastly more complex over a very short period of time. Researchers must now deal with fiendishly difficult multidisciplinary and technological demands. They are plagued with more and more bureaucracy, suffer constant interruptions by all kinds of urgencies, and must keep ever-longer schedules that increasingly lead to less efficient work. Those stressors affect everyone in the sciences, from the group leader to the newcomer.

I recommend Tantra's book to all readers of PHYSICS TODAY. Group leaders will gain perspective on the difficulties novice researchers face and will perhaps also gain some empathy toward their junior colleagues. Graduate students will benefit from tips that could help them avoid painful learning from experience. I also hope the administrators responsible for so many aspects of research management and the well-being of young researchers will view *A Survival Guide for Research Scientists* as required reading.

Pedro C. Marijuán
Zaragoza, Spain

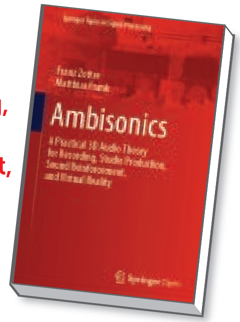
ISTOCK.COM/-OXFORD-

A late-19th-century Edison phonograph.

Re-creating the physical experience of sound

Ambisonics
A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality

Franz Zotter and Matthias Frank
Springer, 2019. $59.99



The recording and reproduction of sound has long been a source of fascination for scientists and engineers. The phonograph, invented by Thomas Edison in 1877, was arguably the first device that could both record and reproduce an acoustic signal. Although it was one of the most remarkable inventions of its time, the phonograph did not attempt to convey any spatial characteristics of the recorded sound field; it simply recorded sound and replicated the signal through a single acoustic source. The monophonic sound field the phonograph generated could not reproduce the original sound's spatial variability.

Over the next several decades, researchers made various attempts to replicate the spatial characteristics of a recorded sound field, without much practical progress. In the 1930s, however, Alan Blumlein invented stereo sound. One technique involved recording a sound field with two microphones, one with sensitivity to sound waves from all directions and one with a figure-eight directivity pattern. When the signals from the two microphones are played back over a pair of loudspeakers spaced carefully apart, a centrally located listener experiences, at least to some extent, the illusion of directional sound.

The invention of ambisonics in the 1970s by Michael Gerzon, Peter Fellgett, and Peter Craven extended Blumlein's technique. As Franz Zotter and Matthias Frank explain in the opening pages of *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, first-order ambisonics allows a recording studio to use four coincident microphones. One microphone is uniformly sensitive and three use figure-eight directivity patterns aligned to the x-, y-, and z-axes of a Cartesian coordinate system. Appropriate processing of those four microphone signals, along with a six-loudspeaker playback system, yields an approximate reconstruction of the directions of arrival of the recorded sound.

The book's first chapter concisely describes those microphone techniques and related approaches and provides the reader with a solid framework for understanding the basic concepts behind ambisonics. Chapter 2 covers numerous experiments that capture how well listeners perceive a change in the direction of arrival of sound as the amplitudes of the inputs to the loudspeakers are varied, or "panned" in the terminology of acoustics. Ville Pulkki's vector-base amplitude panning (VBAP) technique is the subject of chapter 3. It is a straightforward and successful approach to determining the amplitudes of the inputs to arbitrarily arranged loudspeakers in order to generate the illusion of sound coming from a location between the loudspeakers.

The meatiest material, higher order ambisonics, is covered in chapter 4. Zotter and Frank introduce the reader to the spherical harmonic decomposition of the sound field in order to determine the loudspeaker inputs. They also explore the relationship of VBAP to higher-order ambisonics and describe various refinements that can improve the listener's experience of a sound recording. Subsequent chapters deal with signal flow effects, ambisonic microphone arrays, and compact loudspeaker arrays.

Zotter and Frank include an extremely useful bibliography of research in the field and provide many practical and free software options. The authors also helpfully describe several experiments of what listeners perceive as the source of a sound generated by the various recording techniques and their associated panning functions. However, they barely discuss the extent to which various recording strategies are able to replicate the physical properties of the recorded sound field, particularly in the earlier chapters. Chapter 6, on higher-order ambisonic microphones, comes closest to providing some physical insight; it presents the classical problem of a rigid sphere scattering sound waves, shows the steps necessary to reproduce a sampled version of the sound field, and offers some helpful simulations of the resulting pressure distributions.

Ambisonics makes some useful contributions, but the picture is far from complete. There is still room to provide an even deeper understanding of those approaches to sound recording and reproduction. The subject will doubtless continue to fascinate scientists and engineers for some years to come.

Philip Nelson
University of Southampton
Southampton, UK

The Nebra Sky Disk, crafted around 1600 BC.

The rich past of astronomical discovery

From Cave Art to Hubble
A History of Astronomical Record Keeping

Jonathan Powell
Springer, 2019.
$29.99 (paper)

Exploring the history of astronomy is a more challenging journey than one might expect. To understand how humans have viewed the stars, readers must be ready to grapple not only with astronomical concepts but with archaeological discoveries, ancient mythology, and the human imagination. South Wales astronomer and author Jonathan Powell daringly navigates those obstacles for us in his *From Cave Art to Hubble: A History of Astronomical Record Keeping*.

Powell opens *From Cave Art to Hubble* with the black hole in the Messier 87 galaxy, a staple of black hole research for more than two decades. The author then turns his attention to archaeology and tells us about a collaboration between historian of religion Alistair Coombs and chemical engineer Martin Sweatman. The two argue that images of animals and a human–bird hybrid at the French Lascaux cave from around 15 000 BC and engravings on T-shaped monoliths at Göbekli Tepe in Turkey from roughly 9000 BC memorialize comets and terrestrial impacts. Although the paintings and monuments contain no obvious images of comets and no evidence exists of terrestrial impacts during the time of the Lascaux cave paintings, new geological evidence supports the idea that a comet impact may have occurred around the time of the Göbekli Tepe site's construction.

The geological evidence is intriguing, but we will never be able to recapture the thoughts of Paleolithic artists to learn if some of their greatest works were inspired by astronomical catastrophes or by beacons of light in the night sky. Powell might have done well to consult two other resources on the origins of astronomy: *The Roots of Civilization: The Cognitive Beginnings of Man's First Art, Symbol, and Notation* (1971) by the late Peabody Museum scholar Alexander Marshack and the documentary *Lascaux, le ciel des premiers hommes* (*Lascaux, the sky of the first men*; 2007) depicting the work of archaeoastronomer Chantal Jègues-Wolkiewiez. Marshack meaningfully demonstrated that prehistoric humans etched and painted lunar and seasonal time-factored markings on cave walls and portable objects during the Upper Paleolithic. Jègues-Wolkiewiez's work argues that the cave art at Lascaux depicts details of specific constellations as they were seen in the night sky during that epoch.

Powell regains his footing when he moves from Lascaux and Göbekli Tepe to a host of other astronomical artifacts: the Nebra Sky Disk of the German Bronze Age, the Dendera Zodiac from ancient Egypt, and the Roman Farnese Atlas. Although those artifacts have been thoughtfully analyzed by scholars in a range of disciplines, they are not often discussed in books about astronomy or the history of science, which makes them a valuable addition to Powell's work. In those passages, *From Cave Art to Hubble* convincingly ties the ancient past to the present by introducing readers to artifacts that have been astronomically dated and that we can relate to our own night sky.

Powell then takes the reader to the area of his core expertise. *From Cave Art to Hubble* covers historical records of supernovae from around the world; observations on the movements of planets from ancient times; astronomical timekeeping among the ancient Chinese, Egyptians, and Persians; the global historical observation of comets; and an overview of the great observatories in the past four centuries. The depth of this material reflects the author's clear knowledge of the subject matter.

Powell finishes the book with a discussion of the Younger Dryas impact hypothesis, which suggests that a comet or disintegrating asteroid struck Earth some 12 800 years ago in the Northern Hemisphere. That terrestrial impact is proposed to have dramatically shifted the climate into a harsh winter period that led to the extinction of Pleistocene megafauna, influenced the course of civilization, and may be connected to Göbekli Tepe.

The connections between modern discoveries and ancient astronomical observations allow us to ponder all of the intellects that enabled astronomy to be not just science for the sake of science, but a reflection of our inherited interconnectedness to the cosmos. Powell includes a comprehensive glossary that may be of value to both the specialist astronomer and the general audience. However, the book does not include a bibliography, which would have been helpful for further exploration and for ascertaining the strength of the author's sources.

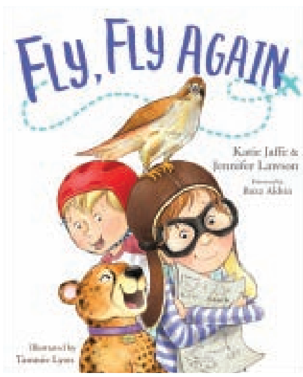
From Cave Art to Hubble is a book to be read and then reread for further reflection as new data and artifacts from our past are uncovered. Powell's work scatters beams of light in the darkness of our astronomical heritage, showing us the rich history of humans exploring the planets, stars, and comets that fill our night sky. The book is a must-read that enables readers to explore both where we have come from and the possibilities that lie ahead.

Bernie Taylor
Portland, Oregon

NEW BOOKS & MEDIA

Fly, Fly Again

Katie Jaffe and Jennifer Lawson, illustrated by Tammie Lyon
Greenleaf Book Group Press, 2020. $15.95



Neighbors Jenny and Jude combine their knowledge of lift and steering to build a flying craft in this colorful book aimed at young elementary school students. Like many children's books, *Fly, Fly Again* is written in rhymes, but their meter often feels a beat or two off and can be awkward to read aloud. But the charming illustrations and the encouraging message about learning and trying again make this a good choice for young tinkerers. Astronaut Buzz Aldrin contributes a short foreword. —MB

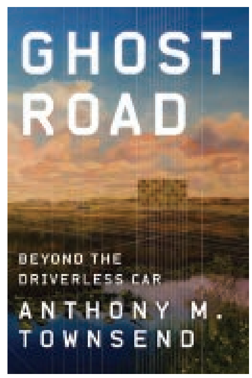
Thinkrolls
Logic Puzzles for Kids

Avokiddo
iTunes, 2019 (version 1.4.3). $3.99

Aimed at children ages 3–8, this physics-based tablet and cellphone game challenges players to solve increasingly difficult mazes by avoiding or overcoming obstacles. Players have to climb, bounce, and roll their way through the puzzles and use basic physics concepts like momentum and force to advance. There is an easy mode for kids under six and a hard mode for older children. The game is also available for Android and Kindle devices. —MB

Ghost Road
Beyond the Driverless Car

Anthony M. Townsend
W. W. Norton, 2020. $27.95



Like the magical flying carpets of Arabian lore, driverless cars could promise a wondrous future of carefree transportation—or they could hasten the decline of civilization as we know it. That uncertain future is the theme of *Ghost Road*, in which Anthony Townsend, an urbanization and digital technology consultant, discusses how automated vehicles (AVs) could transform not only transportation but also daily life, city planning, commerce, and more. Whether AVs lead to safe, energy-efficient vehicles affordable to all or ever-increasing traffic congestion, carbon emissions, and class inequalities, Townsend argues, is entirely up to us and the corporate and public policies we put in place. —CC

The Fab Lab with Crazy Aunt Lindsey

Lindsey Murphy

YouTube and Facebook Live, 2010–present

Science communicator Lindsey Murphy hosts the long-running and beloved YouTube series *The Fab Lab with Crazy Aunt Lindsey*. Murphy encourages her young viewers to get hands-on by making lemon batteries, baking doughnuts, and creating eco-friendly play dough. She also highlights scientists and inventors of color, such as carbon filament inventor Lewis Latimer and Alfred Cralle, who created the first one-handed ice cream scoop. Murphy's newest offering, *Digital Daycare*, is aimed at parents and children under stay-at-home orders during the COVID-19 pandemic. It's a series of two-hour videos in which Murphy guides kids through science and craft projects. —MB

Entangle

Physics and the Artistic Imagination

Ariane Koek, ed.

Hatje Cantz, 2019. $32.00 (paper)

The companion catalog to an art exhibition mounted at Bildmuseet at Sweden's Umeå University, *Entangle* presents the work of 14 contemporary artists who have been inspired by particle physics. Among the themes they incorporated were entropy, gravity, light, matter, space, and time. The pieces span various media, including painting, sculpture, photography, and fashion. In addition to full-page color images of the artworks, the catalog includes scholarly essays by such leading science writers as Philip Ball and Carlo Rovelli and interviews with artists and physicists. Ariane Koek, founder of the Arts at CERN program, served as curator of the exhibition and editor of the catalog. —CC

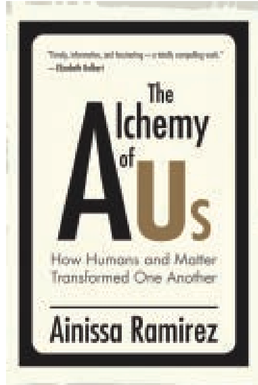
Telling Science Stories

Reporting, Crafting and Editing for Journalists and Scientists

Martin W. Angler

Routledge, 2020. $44.95 (paper)

"Everybody loves stories," writes science journalist, storyteller, and science blog editor Martin Angler. In his latest book, Angler encourages journalists and scientists alike to embrace the use of fiction-writing techniques to convey the latest science discoveries in a way that is both informative and compelling. In addition to chapters on such topics as story selection, narrative structure, language and style, and literary devices, Angler discusses the scientific method and how to locate story elements in scientific papers. A list of review questions, references, and links to online articles and tools rounds out each chapter. —CC



The Alchemy of Us

How Humans and Matter Transformed One Another

Ainissa Ramirez

MIT Press, 2020. $27.95

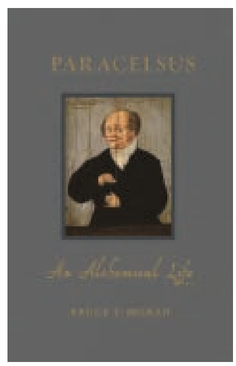
In this readable and entertaining popularization, materials scientist Ainissa Ramirez uses eight inventions—steel rails, photographic film, and silicon chips among them—to show how materials development has been shaped by human interests and how new materials can, in turn, shape society. Ramirez includes both well-known figures such as Thomas Edison and lesser-known scientists. The section on the development of Pyrex, which highlights the work that many women scientists put into creating the bakeware, is especially fascinating. —MB

Paracelsus

An Alchemical Life

Bruce T. Moran

Reaktion Books, 2019. $22.50



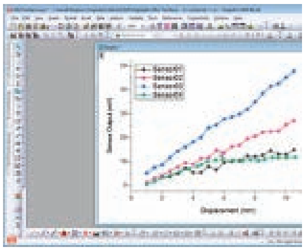
In his latest book, historian of science Bruce Moran focuses on the life of Theophrastus von Hohenheim, better known as Paracelsus, a 16th-century scholar, physician, and alchemist. Although very few of his writings were published during his lifetime, once they were in print, Paracelsus became a major influence on Renaissance medicine. He advocated for the importance of observation in developing new and better medical techniques, but his beliefs were also infused with mysticism and religion. Moran argues that to understand Paracelsus, we must discard modern notions about divisions between magic and science and approach his beliefs on his own terms. —MB PT

NEW PRODUCTS

Focus on test, measurement, software, and instrumentation

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

Andreas Mandelis



Data analysis and graphing software

OriginLab has released version 2020 of its Origin and OriginPro data analysis and graphing software. New features and improvements include mini toolbars for easier 2D graph customization: Buttons in a pop-up provide access to common options for quickly changing graphs without opening complex dialogs. Because the new version makes full use of a processor's multicore architecture, the speed of importing large data files has been improved by a factor of 10 or more compared with previous versions and with Excel 2016. Speed enhancements have also been made in other areas, such as peak analysis, contour plotting, and data import from third-party files. New graph types include density dot and before–after plots, color dots, and dendrograms. New apps include 2D Peak Analyzer, Import NMR Data, and LeCroy Connector. ***OriginLab Corporation***, *One Roundhouse Plaza, Ste 303, Northampton, MA 01060, www.originlab.com*

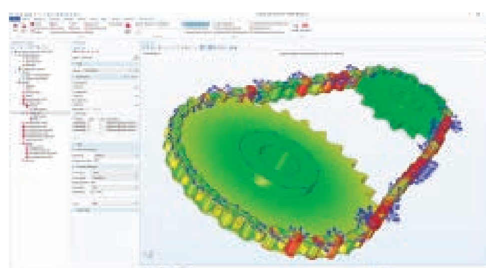
Tabletop system for Hall measurements

The MeasureReady FastHall Station from Lake Shore Cryotronics is a fully integrated tabletop system for performing fast, highly precise Hall effect measurements. The station features Lake Shore's proprietary M91 FastHall measurement controller, a Windows 10 PC, a 1 T permanent magnet, a high-precision sample holder, and the necessary software and cabling. According to the company, the system requires less setup time than similar apparatus. The electronically shielded, low-noise sample space with guarded contacts allows users to make high-quality measurements for quick and easy derivation of carrier type, carrier concentration, mobility, and Hall coefficient properties of a sample. Also available are a gate-bias instrument option and a liquid-nitrogen option for converting the standard room-temperature station to a cryogenically cooled sample space. ***Lake Shore Cryotronics Inc***, *575 McCorkle Blvd, Westerville, OH 43082, www.lakeshore.com*

Atomic force microscope accessories

Oxford Instruments Asylum Research now offers several accessories for its versatile Jupiter XR large-sample atomic force microscope (AFM). They include the PolyHeater for heating samples up to 300 °C; the CoolerHeater, with a cooling and heating range of −30 °C to 120 °C; a fluid cell and probe holder for liquid imaging; and conductive AFM probe holders for nanoelectrical measurements. According to the company, the accessories will broaden experimental possibilities for multiuser, multidisciplinary laboratories. Many research fields require environmental control, like that provided by the PolyHeater and the CoolerHeater, to perform such tasks as mimicking real-life conditions and exploring thermal properties. The new probe holders will facilitate AFM imaging and nanoelectrical measurements, which are essential when working with semiconductor, 2D, and other functional materials. ***Oxford Instruments Asylum Research***, *6310 Hollister Ave, Santa Barbara, CA 93117, www.oxinst.com*

Modeling and simulation software



Comsol has updated its Multiphysics software for creating physics-based models and simulation applications. New in version 5.5 are geometry modeling tools and solver technology for acoustics simulation. The Design Module now provides a sketching tool for easier creation and more versatile parametric control of geometry models. Its users can more easily assign dimensions and constraints to planar drawings for 2D models and 3D work planes. New functionality based on the time-explicit discontinuous Galerkin method enables efficient multicore computations of ultrasound propagation in solids and fluids, including realistic materials featuring damping and anisotropy. The method also has low-frequency applications, such as in seismology. Two new products are the metal-processing and porous media flow modules. ***Comsol Inc***, *100 District Ave, Burlington, MA 01803, www.comsol.com*

Wavelength meter for optical transceiver testing

Bristol Instruments has developed its 338 series optical wavelength meter to improve the efficiency of optical transceiver wavelength testing. With a high measurement rate of 25 Hz, it delivers greater production throughput at lower cost than comparable systems, according to the company. The 338 optical wavelength meter uses Michelson interferometer–based technology with FFT analysis to measure the wavelength of CW or modulated signals to an accuracy of ±1.0 pm. Continuous calibration with a built-in wavelength standard ensures reliable test results. A convenient touch-screen display controls the system and shows the wavelength and power measurements in various formats. The data can also be sent to a PC. ***Bristol Instruments Inc****, 770 Canning Pkwy, Victor, NY 14564, www.bristol-inst.com*

Wideband measurement analysis

Keysight has designed a flexible, economical single-channel instrument to accelerate development of next-generation mmWave communications, satellite communications, and radar applications. The UXR0051AP Infiniium UXR-series oscilloscope offers a frequency range of 110 GHz and a standard analysis bandwidth of 5 GHz. It displays a very low −158 dBm/Hz average noise level from 28 GHz to 85 GHz, which enables error-vector-magnitude measurements of golden receiver quality on low-power wideband signals. The device directly measures wideband signals with up to 10 GHz bandwidth and fundamental frequencies as high as 110 GHz without the need for external downconverters. It can instantly expand to two independently configurable phase-coherent channels for multiple-input and multiple-output measurement support. ***Keysight Technologies Inc****, 1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*

Compact dry vacuum pump

The nXRi high-performance compact dry pump from Edwards offers low ultimate pressure and, according to the company, the highest pumping density on the market. Its compact size, low power, reliability, and low cost make it suitable for a wide range of applications. Initial variants provide pumping speeds of 60 m^3/h and 90 m^3/h, four times as fast as a similar-sized dry pump, the company claims. The nXRi fits easily under a benchtop and, at under 30 kg, is very mobile. It is maintenance-free for up to five years, with no tip-seal or oil change needed. Compared with alternative dry pumps, the nXRi has a 40% smaller footprint, which ensures its seamless integration into analytical instruments and vacuum systems. It is suitable for mass spectrometry, electron microscopy, and leak detection. ***Edwards Ltd****, Innovation Dr, Burgess Hill, West Sussex, RH15 9TW, UK, www.edwardsvacuum.com*

Source meter without need for pulse tuning

The 2601B-Pulse System SourceMeter instrument from Tektronix integrates into one instrument a high-speed current pulser with DC source and measurement functions. It incorporates new PulseMeter technology—an industry first, according to the company—as a source of current pulses as short as 10 μs at 10 A and 10 V without the need to manually tune the output to match device impedance up to 3 μH. That is critical for minimizing device self-heating, which for optical instruments can result in measurement errors and equipment damage. Built-in dual 1 MS/s, 18-bit digitizers enhance the pulser's measurement function, which allows users to acquire pulse current and voltage waveforms simultaneously. The 2601B-Pulse System SourceMeter is suitable for such applications as characterizing semiconductor devices and testing vertical-cavity surface-emitting lasers and LEDs, fault power management, and surge protection. ***Tektronix Inc***, *14150 SW Karl Braun Dr, PO Box 500, Beaverton, OR 97077, www.tek.com*

Smart-connected turbomolecular pumps

Agilent has added two compact turbomolecular pumps to its TwisTorr family of devices: the TwisTorr 305 FS and the TwisTorr 305 IC. Both offer smart connectivity, a new feature for Agilent turbomolecular pumps. When installed on Apple or Android phones, the Vacuum Link app lets users communicate remotely. By typing commands and modifying parameters, users can quickly and easily control the pumps. An advanced function allows users to extract log files to share pump-operating data, which saves time. The model 305 FS, which features floating suspension, is a stand-alone unit with an external remote controller; the model 305 IC has an integrated controller. The pumps' small footprint makes them suitable for OEMs and other companies that want to integrate the pump in an instrument. ***Agilent Technologies Inc***, *5301 Stevens Creek Blvd, Santa Clara, CA 95051, www.agilent.com*

Analog and digital signal acquisition

Spectrum Instrumentation has extended the capabilities of its LXI/Ethernet digitizers by making a mixed-mode testing option available on four models. All feature eight analog input channels that synchronously sample signals at rates up to 5, 20, 40, or 125 MS/s with 16-bit resolution. The DN2.59x-08-Dig option adds eight digital lines to the three multipurpose input/output lines that come as standard with the units. The enhanced digitizers simultaneously acquire 8 analog and 11 digital signals in a fully synchronous fashion, and front-panel BNC connectors provide easy access to all 19 channels. Software drivers allow customized setups that can generate perfectly matching mixed-mode solutions. The units come with the tools needed to integrate them into virtually any test system, including mechatronics, vibrational studies, and control systems. ***Spectrum Instrumentation Corp***, *401 Hackensack Ave, 4th Flr, Hackensack, NJ 07601, https://spectrum-instrumentation.com*

High-resolution multichannel event timers

PicoQuant's MultiHarp 150 high-throughput multichannel event timers feature the company's latest time-correlated single-photon-counting electronics for fast, high-resolution fluorescence lifetime imaging and multichannel photon correlation. The latest MultiHarp 150 models—the 4P, 8P, and 16P—have 4, 8, or 16 detection channels and offer improved timing precision with 10 ps minimum bin width and jitter better than 45 ps (rms). According to the company, that is the best time resolution of any currently available event timer that has subnanosecond dead time. The MultiHarp 150 tabletop units are versatile and easy to use. They have a USB 3.0 interface and are suitable for many time-resolved applications in the life and materials sciences, metrology, and single-photon-based quantum technologies. ***PicoQuant***, *Rudower Chaussee 29, 12489 Berlin, Germany, www.picoquant.com* PT

OBITUARIES

Philip Warren Anderson

Philip Warren Anderson, one of the intellectual giants who shaped and nurtured the rapid growth of condensed-matter physics during the second half of the 20th century, died on 29 March 2020 in Princeton, New Jersey. He made fundamental contributions to diverse subfields, including antiferromagnetism, superexchange, dirty superconductors, the x-ray singularity problem, localization, superfluidity in helium-3, spin glasses, quantum spin liquids, local moments in metals, poor-man's renormalization, and cuprate superconductivity. Many of those concepts now carry his name. He was a corecipient, along with Nevill Mott and John Van Vleck, of the 1977 Nobel Prize in Physics for "fundamental theoretical investigations of the electronic structure of magnetic and disordered systems."

Anderson was born on 13 December 1923 in Urbana, Illinois. After a stint at the Naval Research Laboratory during World War II, he obtained his PhD in 1949 from Harvard University, working under Van Vleck. In 1949 he joined Bell Labs and its group of talented physicists, which included Bernd Matthias, Peter Wolff, Robert Shulman, William Shockley, and Charles Kittel. Their strong influence on the company to invest in basic research had a great effect on the labs for the rest of the century. From 1967 to 1975, Anderson worked part time at Cambridge University before joining the faculty of Princeton University. In 1984, after retiring from Bell Labs, he started as a full-time professor and became emeritus in 1996.

Anderson is perhaps best known for his 1958 prediction that sufficiently strong disorder can turn metals into insulators via a process now known as Anderson localization. Before his work, the common view was that electron waves are extended throughout the material. Anderson showed that at low temperatures, disorder can cause the waves to be localized in space and stop conducting current.

In the 1950s and 1960s, Anderson elucidated how a combination of quantum mechanics and strong repulsion between electrons causes electron spins to form local moments; his insight laid the foundation of the modern theory of magnetism. After John Bardeen, Leon Cooper, and J. Robert Schrieffer proposed their pairing theory of superconductivity in 1957, Anderson became a major contributor to the topic. He was the first to recognize the importance of the phase of the superconducting wavefunction and how it is quantum mechanically conjugate to the number of Cooper pairs.

In principle, phase fluctuations lead to a collective mode of gapless excitations, an example of Goldstone's theorem; such excitations are observed in neutral superfluids but not in superconductors. Anderson realized that the coupling of Cooper pairs to the electromagnetic field boosts the mode to a finite frequency, where it merges with the plasma mode. He proposed in 1962 that the mechanism removes the roadblock—namely, the problem of unwanted massless Goldstone particles—facing unified field theories based on broken symmetry. Two years later Peter Higgs and others reached the same conclusion via more formal routes. As Higgs wrote in his Nobel lecture, "The Goldstone massless mode became the longitudinal polarization of a massive spin-1 'photon,' just as Anderson had suggested." The Anderson–Higgs mechanism is now a cornerstone of both particle and condensed-matter physics.

A year after the 1986 discovery of high-temperature superconductivity in cuprates, Anderson published an enormously influential paper in *Science* pointing out that the key physics is the introduction of charge carriers ("holes") into the insulating state that arises from strong electron–electron repulsion. He recalled a 1973 paper that introduced the notion of quantum spin liquids, in which magnetic moments fail to achieve long-range order because of quantum fluctuation and instead form a state that he dubbed a "resonating valence bond" (RVB). He proposed that in a cuprate, when holes are introduced into that state, it becomes a superconductor.

Those revolutionary ideas met stiff resistance from the community. Although the specific mechanism he proposed for superconductivity remains controversial, many of the ideas he introduced in the 1987 paper, including the notion that superconductivity is a favorable ground state in a strongly repulsive system, have gained wide acceptance. The RVB state is the archetypal example of a quantum spin liquid, currently a vigorous area of research.

Anderson also suggested that the ex-

Philip Warren Anderson

AIP EMILIO SEGRÈ VISUAL ARCHIVES/PHYSICS TODAY COLLECTION

citations of a quantum spin liquid behave as electrons that have lost their charge but retain their spin. That early example of "fractionalization" has found support both in exactly soluble models and in real materials. Time will tell, but Anderson's spin-liquid work may well be remembered as his most profound and prescient.

In a 1972 article entitled "More is different," Anderson outlined the antireductionist view that each layer of nature is as worthy of fundamental investigation as the most microscopic ones. Those laws cannot anticipate, much less explain, the rich variety of macroscopic systems' fascinating complex behavior, such as superconductivity, chaos, and emergent phenomena. That view has deeply influenced condensed-matter physics and other areas of science.

In addition to the Nobel Prize, Anderson was awarded the American Physical Society's Oliver E. Buckley Prize in 1964 and the National Medal of Science in 1982. He had a lifelong interest in the game of Go dating from a yearlong visit to Japan in 1953–54, and he attained the rank of first-Dan master. In 2007 the Nihon Ki-in, Japan's association for Go, gave him a lifetime achievement award.

William F. Brinkman
N. Phuan Ong
Princeton University
Princeton, New Jersey
Patrick A. Lee
Massachusetts Institute of Technology
Cambridge

Alvin Virgil Tollestrup

Alvin Virgil Tollestrup, a key figure in the development of Fermilab, notably the creation of the Tevatron, succumbed to cancer on 9 February 2020 in Warrenville, Illinois.

Born on 22 March 1924 in Los Angeles, Alvin moved with his family when he was six to Logan, Utah, where his grandfather's position as a psychology professor helped them weather the Depression. There young Alvin had a basement laboratory with a chemistry set and electronic components.

Later the family moved to Salt Lake City. In his senior year of high school, while preparing a term paper for an inspiring chemistry teacher, Alvin devoured Robert Millikan's 1935 book on cosmic rays, wrote about the latitude effect, and began to dream of Caltech and a career in research.

Alvin graduated in 1944 from the University of Utah with a degree in general engineering; he had enrolled in physics courses rather than less interesting ones that a specific engineering degree would require. He qualified for the US Navy's radar school, where he could learn about microwave and pulse techniques. He was assigned to install and test equipment and, eventually, to teach radar technology.

Supported by the GI Bill, Alvin went to graduate school at Caltech in 1946 and worked with William Fowler and Charles Lauritsen in the Kellogg Radiation Laboratory. Their group developed a new method to determine masses and mass defects of nuclides by precisely measuring the energy released in nuclear reactions. Its application to the light elements became the subject of Alvin's 1950 dissertation.

Continuing on at Caltech as a research fellow, Alvin began a move toward high-energy physics with a series of pion photoproduction measurements. He used Caltech's 500 MeV electron synchrotron to map out the 3–3 resonance—now known as $\Delta(1238)$.

Alvin spent 1957–58 at CERN, the fledgling European laboratory, on an NSF fellowship. He worked with Giuseppe Fidecaro and others on the 600 MeV synchrocyclotron and reported the first detection of the rare decay of a charged pion into an electron and a neutrino. The first experiment at CERN's first accelerator, it supported the nascent $V - A$ description of weak interactions. On his return to Caltech, Alvin was promoted to associate professor, and he became a full professor in 1962.

Alvin went to Fermilab on sabbatical in April 1975, when a magnet-development program had been launched to realize Robert Wilson's vision of a superconducting synchrotron. The early model magnets were primitive, but Alvin brought a focus on underlying physical principles and a faith that judicious application of freshman physics could master any challenge. He regarded failures as discoveries to be made. His approach is exemplified in his 1979 report *The Amateur Magnet Builder's Handbook*. One of his indispensable innovations was to secure the superconducting magnet coils with interlocking stainless steel collars. For his contributions to the Tevatron, which was for two decades the world's highest-energy accelerator, Alvin received the National Medal of Technology in 1989.

It would be several years before the Tevatron—as a proton accelerator and proton–antiproton collider—became an official project, but Alvin was already looking forward to the research it would enable. In December 1976 Fermilab created a colliding-beams department. Alvin was appointed to a five-person steering group that guided the exploration of both detector and accelerator issues, including the production and cooling of antiprotons. In less than a year, Alvin became department head.

By the spring of 1979, the first design report was produced for a detector and superconducting solenoid magnet that would become the Collider Detector at Fermilab (CDF). Alvin was central to expanding the collaboration into an international team that included institutions in Japan and Italy. The CDF group submitted the formal design report to the Department of Energy in the summer of 1981, and first collisions were seen in the detector four years later. During the exciting period of the top-quark discovery in 1994 and 1995, Alvin held regular sessions with young CDF physicists, especially the postdocs in the Fermilab group.

Alvin was always on the lookout for unique and promising physics opportunities. For more than two decades, he worked with an international group of accelerator and detector physicists to address the challenges of designing a high-energy collider that would use short-lived muons rather than stable electrons or protons. He dove into the technologies required for high-field accelerator magnets and the operation of RF cavities in high magnetic fields.


Alvin Virgil Tollestrup

FERMILAB CREATIVE SERVICES

As a nonagenarian, Alvin attached himself to a group of young physicists exploring novel methods for detecting the hypothetical axion. He drew on his accumulated radar wisdom and what he liked to call "those damn Smythe problems" from William Smythe's classic *Static and Dynamic Electricity*. In a series of beautiful technical notes, Alvin produced a complete simulation of how nonlinear dielectric crystals interact with microwave cavity modes.

Alvin delighted in identifying promising early-career researchers. He helped them to create detailed research plans and continually challenged them to develop a deep and thorough understanding of their observations. Many young researchers savored the chance to experience firsthand his incisive approach to problem solving. The Tollestrup Award for Postdoctoral Research, presented annually by the Universities Research Association, celebrates Alvin as talent scout and mentor. We cherish the memory of his achievements and his example.

Chris Quigg
Fermi National Accelerator Laboratory
Batavia, Illinois
Mel Shochet
University of Chicago
Chicago, Illinois PT

INDEX TO ADVERTISERS

ADVERTISING SALES MANAGEMENT

American Institute of Physics
One Physics Ellipse
College Park, MD 20740
+1 301.209.3393
www.physicstoday.org/advertising

Sales and Marketing Manager
Christina Unger Ramos • cunger@aip.org

Recruitment Advertising
ptjobs@aip.org

Ad Traffic Coordinator
Unique Carter • ucarter@aip.org

ADVERTISING SALES CONSULTANTS

USA and CANADA

AL, AK, AZ, AR, CO, HI, ID, IL, IN, IA, KS, KY, LA, MI, MN, MS, MO, MT, NE, NV, NM, ND, OH, OK, SD, TN, UT, WV, WI, WY
CEL Associates II
Merrie Lynch
+1 617.594.4943
merrie.lynch@celassociates2.com

CA, CT, DE, FL, GA, ME, MD, MA, NH, NJ, NY, NC, OR, PA, RI, SC, VT, VA, WA, and Canada
CEL Associates II
Jody Estabrook
p: +1 774.283.4528
m: +1 978.835.2013
jodyestabrook@comcast.net

INTERNATIONAL

Northern and Western Europe
John Gold
+44 208.641.7717
johnfgold@gmail.com

Central and Eastern Europe
Leo Ploner
+49 172.8231.062
leo@ploner.de

Asia/Pacific
Jake Wilson
+61(0)413 949 644
jwilson@imraust.com.au

PHYSICS TODAY

www.physicstoday.org/advertising

Aaron Slepkov is a Canada Research Chair in the physics of biomaterials at Trent University in Peterborough, Ontario, Canada.

Fruit photonics and the shape of water

Aaron Slepkov

To microwaves, grapes are resonant, spherical blobs of water.

Cut a grape in half, leaving the hemispheres attached by an isthmus of skin. Then irradiate the pair in a household microwave oven. Those are the directions for a deceptively simple experiment Patrick Michaud published online in 1994. Within a few seconds, sparks emanate from the skin bridge and ignite a plasma. For decades the popular parlor trick has delighted and surprised millions of YouTube viewers, science-fair participants, and others, including readers of this magazine (see Back Scatter, PHYSICS TODAY, October 2017, page 96). In my lab at Trent University the oddity has blossomed into a research project involving several undergraduates, many pounds of charred fruit, and a dozen broken microwave ovens.

This past year we published our own account of the phenomenon. To some people the surprise comes from seeing biological tissue spark in ways they have come to expect only from metallic objects. But to us and other physicists, the surprise comes from the deeply subwavelength nature of electromagnetic-energy concentration. Microwave ovens operate at a frequency of 2.5 GHz, which corresponds to a free-space wavelength of 12 cm. With diameters ranging from 1 to 2 cm, grapes are much smaller than microwave wavelengths. Charred markings on the grapes are smaller still, on the order of millimeters. That size, mysteriously, is about one-hundredth of the 12 cm wavelength.

Two related photonic mechanisms can create intense, highly localized electromagnetic-field "hot spots:" surface plasmon resonances (SPRs) in metals, and morphology-dependent resonances (MDRs) in nonmetals. The SPRs are resonant surface-charge oscillations induced at the nanoscale (see the article by Mark Stockman, PHYSICS TODAY, February 2011, page 39). By analogy to atomic resonances that combine, or hybridize, to form new molecular resonances, SPR resonances can hybridize to produce super-intense hot spots at the nexus of dimers and larger clusters. Such hot spots are driving applications across a range of fields, such as chemical sensing, single-molecule spectroscopy, and photodynamic therapy.

The MDRs in transparent dielectric particles can likewise display optical resonances that yield electromagnetic hot spots. Recent research shows that when the nonmetal particles have a sufficiently large index of refraction, they can mimic the near-field hot spots of metallic SPRs. But unlike metals, transparent dielectrics admit electric fields into their interior, which makes internal light modes an important characteristic of MDRs.

A central question in our sparking-grape research is which of the two mechanisms is responsible for creating microwave-field concentrations intense enough to ignite a plasma. At first glance, the ion-laden cut grapes appear sufficiently conductive to make a plasmonic explanation plausible. And I initially wondered whether the hemispheres were acting like short metallic antennas. Further investigations, however, found that the answer lies in the grapes' behavior as dielectric spheres, not metallic ones. In this Quick Study, I describe why treating grapes as simple balls of water is the key to explaining why they spark in the microwave oven and why MDRs are the source of the phenomenon.

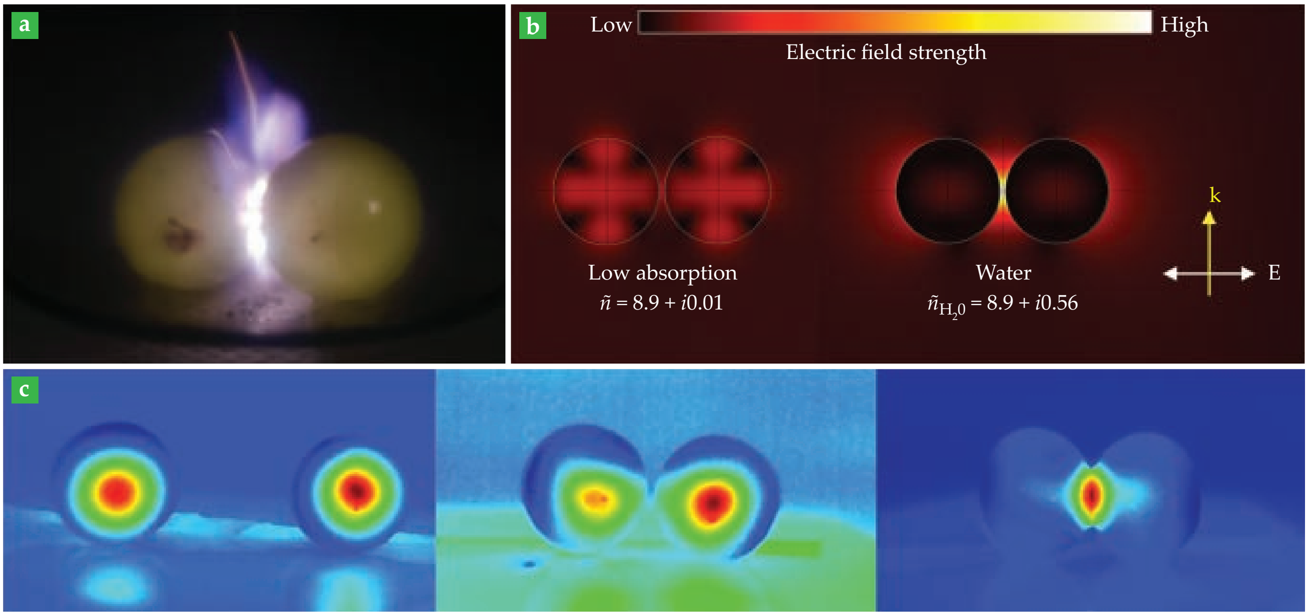
Spheres of water

The first hint that something was amiss in the plasmonic explanation was our discovery that grapes do not need to be cut in half to generate plasma. As long as they are at least near contact, a pair of whole grapes will spark, as shown in the figure's panel a. The skin bridge is merely a convenient way to keep the objects close together. That finding led us to a breakthrough hypothesis—that the biology is irrelevant, and that grapes simply act like blobs of water in air. To confirm the hypothesis, my group made hydrogel balls, small sodium polyacrylate beads that hydrate to form skinless grape-sized (or larger) balls containing more than 99.6% pure water. Much like grapes, the beads never sparked when they were alone, but routinely created a microwave plasma when they were dimers in contact.

What's so special about water? Transparent over the broad, visible spectrum, and with an index of refraction of 1.3 at visible wavelengths, water is a mundane optical material; it absorbs little radiation and has a refractive index lower than that of glass. However, in the microwave regime, water is an exciting material whose index of refraction is of order 10. That's high enough for water blobs to behave like resonant cavities that strongly confine the microwaves.

The lowest-order MDRs, namely the electric and magnetic dipolar modes, form when the wavelength of incident light inside a dielectric is about the size of the particle's diameter. At 2.5 GHz one would expect to see fundamental optical scattering modes that mimic SPRs in spheres of water 1.3 cm in diameter—the size of grapes. Computer simulations, shown in the figure's panel b, reveal those field-concentration patterns in such spheres. When the absorption of water is included in the simulations, the resonance modes are more uniform and localized near the center. As two grapes are brought together, the modes overlap, merging into the tight space between the grapes, where the electric field grows enormous.

Although impressive, the plasma itself is of little scientific interest. Its presence simply indicates the formation of an intense electromagnetic hot spot. The concentration of microwaves is what prompts our interest. As described above, a



SUBWAVELENGTH ELECTRIC-FIELD CONCENTRATIONS in dielectric dimers. A pair of grapes **(a)** spark in a household microwave oven because of an optical hot spot at their point of contact. **(b)** Electromagnetic simulations in nonabsorbing grape-sized dimers show concentration patterns (left) inside and near their point of contact. The complex refractive index $\tilde{n}$ includes the refractive index and the absorption of radiation. When water's absorption of microwaves is included in the simulation (right), the internal mode structure is far more uniform and the evanescent hot spot outside the absorbing particles dominates. That hot spot creates the grape plasma. **(c)** A thermal-imaging sequence in microwave-irradiated grape hemispheres reveals the mixing of isolated optical resonances as two grapes are brought together. (Images adapted from H. K. Khattak, P. Bianucci, A. D. Slepkov, *Proc. Natl. Acad. Sci. USA* **116**, 4000, 2019.)

key difference between plasmonic resonances and MDRs is the internal optical modes of dielectrics. A measurement of electric-field distributions both inside the dielectric spheres and near their surfaces would thus be strong evidence tying MDRs to the creation of a dimer hot spot.

But measuring such fields directly is difficult, mainly because the hot spots are so small—about one-hundredth the microwaves' wavelength—and easily perturbed by contact probes. Instead, we exploited the fact that aqueous objects absorb at microwave frequencies. We used thermal imaging as an indirect way to measure time-integrated field intensities. Guided by additional three-dimensional finite-element method simulations, thermal imaging provides key evidence of the microwave MDRs in water spheres and of the mixing of their modes, shown in the figure's panel c.

Thermal imaging of irradiated hydrogel and grape hemispheres of various sizes reveals the increasing complexity of those internal resonances. Particles larger than grapes accommodate higher-order modes. Even so, the presence of internal modes does not by itself preclude a surface-conductivity effect. Our best evidence against surface conductivity comes from a fun experiment with quail eggs: Using thermally activated paper, which turns black above 90 °C, we confirmed that a pair of unmodified eggs (about 24 mm in diameter) develops a literal hot spot at the point of contact. When eggs are emptied of their contents, the hot spot disappears. And when they are refilled with water, it reappears.

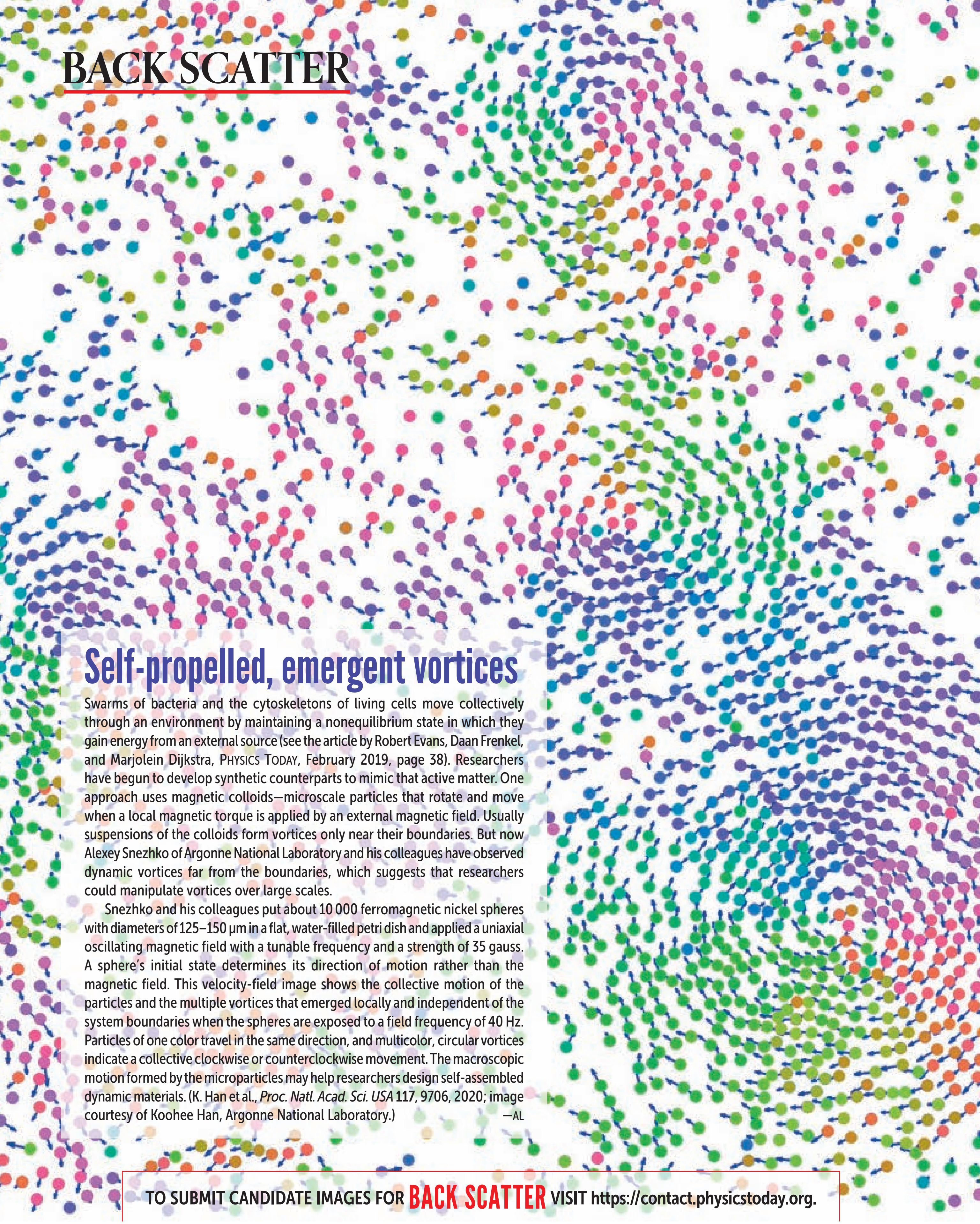
Grape Expectations

Beyond the pyrotechnics, our studies have opened the door to other interesting avenues of research. Early on, we noticed that the dimer tends to vibrate rapidly when irradiated, often just before a plasma ignites. We recently tied that motion to the vaporization of water from the superheated hydrogel surface. Analogous to the Leidenfrost effect in liquid–solid interfaces (see PHYSICS TODAY, November 2018, page 14), a volatile elastic solid can convert thermal energy to mechanical motion. The microwave optical resonance creates a dynamic hot spot that explosively vaporizes the objects at their point of contact and allows them to push off each other.

Such remote activation may find applications in soft robotics. Other, more fanciful applications, such as omnidirectional antennas or MDR-based analogues of surface plasmon lasers, are also possible. In broader terms, the work demonstrates how microwave-water photonic research can be a powerful experimental sandbox for scaled-up investigations of resonant-scattering phenomena that cannot be elucidated at the nanoscale.

Additional resources

▶ H. K. Khattak, P. Bianucci, A. D. Slepkov, "Linking plasma formation in grapes to microwave resonances of aqueous dimers," *Proc. Natl. Acad. Sci. USA* **116**, 4000 (2019); for a video, see https://www.youtube.com/watch?v=wA4uZGRENas.
▶ A. I. Kuznetsov et al., "Optically resonant dielectric nanostructures," *Science* **354**, aag2472 (2016).
▶ H. K. Khattak, S. R. Waitukaitis, A. D. Slepkov, "Microwave induced mechanical activation of hydrogel dimers," *Soft Matter* **15**, 5804 (2019). PT



Self-propelled, emergent vortices

Swarms of bacteria and the cytoskeletons of living cells move collectively through an environment by maintaining a nonequilibrium state in which they gain energy from an external source (see the article by Robert Evans, Daan Frenkel, and Marjolein Dijkstra, Physics Today, February 2019, page 38). Researchers have begun to develop synthetic counterparts to mimic that active matter. One approach uses magnetic colloids—microscale particles that rotate and move when a local magnetic torque is applied by an external magnetic field. Usually suspensions of the colloids form vortices only near their boundaries. But now Alexey Snezhko of Argonne National Laboratory and his colleagues have observed dynamic vortices far from the boundaries, which suggests that researchers could manipulate vortices over large scales.

Snezhko and his colleagues put about 10 000 ferromagnetic nickel spheres with diameters of 125–150 μm in a flat, water-filled petri dish and applied a uniaxial oscillating magnetic field with a tunable frequency and a strength of 35 gauss. A sphere's initial state determines its direction of motion rather than the magnetic field. This velocity-field image shows the collective motion of the particles and the multiple vortices that emerged locally and independent of the system boundaries when the spheres are exposed to a field frequency of 40 Hz. Particles of one color travel in the same direction, and multicolor, circular vortices indicate a collective clockwise or counterclockwise movement. The macroscopic motion formed by the microparticles may help researchers design self-assembled dynamic materials. (K. Han et al., *Proc. Natl. Acad. Sci. USA* **117**, 9706, 2020; image courtesy of Koohee Han, Argonne National Laboratory.) —AL

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