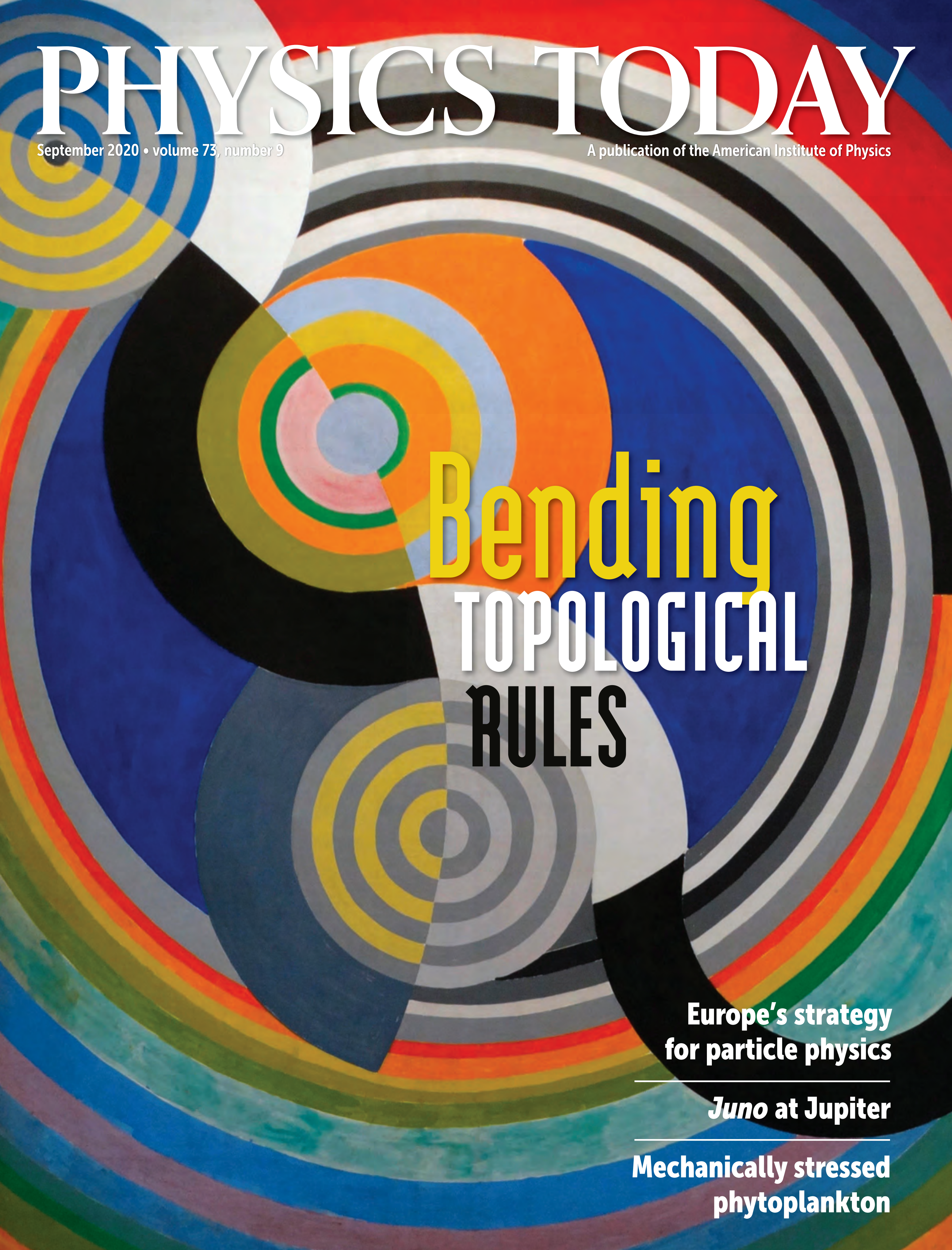
PHYSICS TODAY

September 2020 • volume 73, number 9

A publication of the American Institute of Physics



Bending TOPOLOGICAL RULES

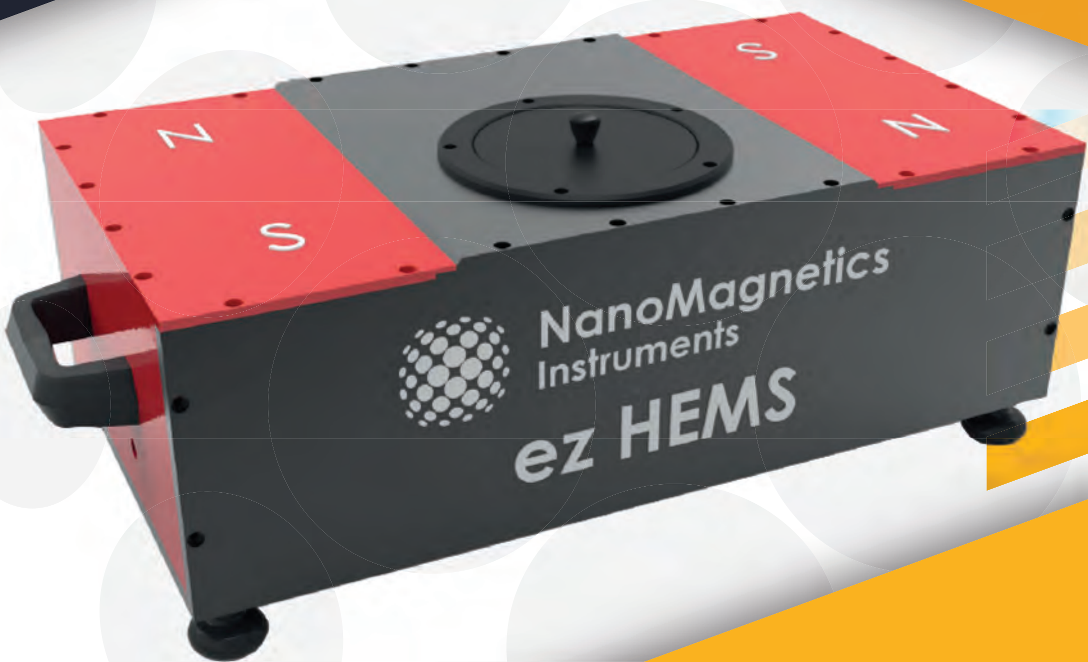
Europe's strategy for particle physics

Juno at Jupiter

Mechanically stressed phytoplankton

NanoMagnetics Instruments

ezHEMS



Technical Specifications

- Characterize / measure the sample resistance, resistivity, Hall coefficient, Hall mobility, carrier concentration or current-voltage characteristics
- 0.6T or 1.0T magnetic field options
- Compatible with Hall bar and van der Pauw shaped samples
- Temperature range 80K-800K
- Sample sizes up to 15mm x 15mm x 3mm

The ezHEMS Measurement System Software Capabilities

sales@nanomagnetics-inst.com
www.nanomagnetics-inst.com

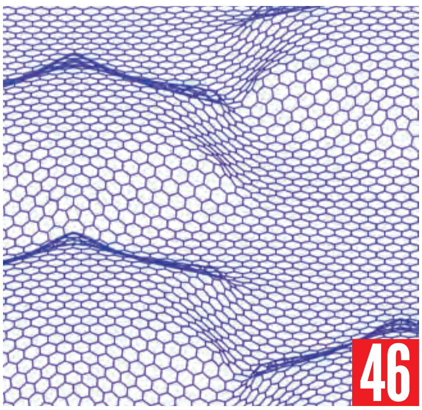
+44 7906 159 508

Suite 290, 266 Banbury Road
Oxford OX2 7DL, United Kingdom

NMInstruments

PHYSICS TODAY

September 2020 | volume 73 number 9



FEATURES

30 Dawn of the topological age?

Arthur P. Ramirez and Brian Skinner

Nontrivial electron band structures may enable a new generation of functional materials.

38 Finnish–Soviet nuclear icebreakers

Saara Matala

During the Cold War, Eastern and Western manufacturers found good reasons to collaborate, even on a technology as sensitive as nuclear vessels.

46 Graphene gets bent

Bastien F. Grosso and Eugene J. Mele

Two-dimensional nanomaterials are bending the rules of the papercraft known as kirigami.

ON THE COVER: In his abstract geometric works, French painter Robert Delaunay (1885–1941) conveyed a sense of constant motion through intersecting disks and rhythmic, flowing bands of contrasting colors. As described in the article by Arthur Ramirez and Brian Skinner on **page 30**, similar motifs, including winding trajectories and intersecting bands of electrons, underpin the nature and potential promise of topological materials. And to learn how periodically modulating a topological system adds a new twist, turn to **page 14**. (Detail from *Rythme n°2, décoration pour le Salon des Tuileries*, 1938, Robert Delaunay. Image from Photo 12/Alamy Stock Photo.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org

WHITE HOUSE

▶ ADA at 30

This summer was the 30th anniversary of the signing of the Americans with Disabilities Act. Although there is evidence of progress, science students and researchers with disabilities report having to constantly advocate for accommodations and equal access to information and spaces.
physicstoday.org/Sep2020a

AWI/MICHAEL GUTSCHE, CC BY 4.0

▶ Icebound

For nearly a year, an icebreaker ship has floated along with the Arctic Ocean ice pack and allowed scientists to take surface, atmospheric, and undersea measurements that will become baseline data for future studies. Three participants describe the unique Mosaic project and the impact of the pandemic.
physicstoday.org/Sep2020b

NTI

▶ Nuclear security

Improving the security of weapons-usable nuclear materials worldwide has slowed in the past two years, according to a recent report by the nongovernmental Nuclear Threat Initiative. (Its CEO, Ernest Moniz, is pictured.) PHYSICS TODAY's David Kramer outlines some of the vital gaps in security.
physicstoday.org/Sep2020c

PHYSICS TODAY (ISSN 0031-9228, coden PHTOAD) volume 73, number 9. 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

8 From the editor

10 Readers' forum

Letters

14 Search & discovery

Cold-atom lattice bends topological rules • An Alaskan volcano, climate change, and the history of ancient Rome • Mechanically stressed phytoplankton light up

22 Issues & events

Cats and llamas could offer a path to coronavirus therapies • Particle physicists hash out long-term strategy for Europe

53 Books

Energy governed by the laws of physics — *Julia Mikhailova* • Modeling the future of Earth's climate — *Peter R. Gent* • A global history of changing knowledge — *Joseph D. Martin* • New books & media

58 New products

Focus on test, measurement, and analytical equipment

60 Obituaries

Eleanor Margaret Burbidge • George Trilling

62 Quick study

Juno at Jupiter — *David Stevenson*

64 Back scatter

Predicting large solar flares

FROM THE EDITOR

Endless front matter

Charles Day

In November 1944, as Allied armies neared Nazi Germany's industrial heartland, President Franklin Roosevelt wrote to electrical engineer Vannevar Bush to request a report on how science could be harnessed in peacetime. "New frontiers of the mind are before us," wrote the president, "and if they are pioneered with the same vision, boldness, and drive with which we have waged this war, we can create a fuller and more fruitful employment and a fuller and more fruitful life."

Bush was an apt choice for author. In June 1941 Roosevelt had appointed him director of a new agency, the Office of Scientific Research and Development, which successfully coordinated scientific R&D for military ends. Bush finished his report eight months after getting the request and delivered it to Roosevelt's successor Harry Truman.

The report, *Science—The Endless Frontier*, advocated establishing a national science foundation to fund curiosity-driven basic research conducted largely at universities.[1] The creation of NSF five years later was not inevitable. Bush and his allies had to prevail over rival proposals for a more applied, mission-driven science agency. As historian Emily Gibson observed recently, the enduring legacy of *Science—The Endless Frontier* lies not just in NSF's continued success but also in the continued debate among policymakers on the extent to which government-funded science should be curiosity driven or mission driven.[2]

Bush's 75-year-old report has another lesson for us. Its literary style and structural organization are exemplary. The main body is divided into chapters, which are further divided into sections of clear, continuous prose. Enumerated lists are used sparingly. Whenever Bush needed to emphasize a point at the end of a section, he used italic, which is easier on a reader's eyes than shouty bold.

Another exemplar of presentation and clarity is the 1962 report of the Royal College of Physicians (RCP), *Smoking and Health*, which summarized the evidence gathered by Richard Doll, Austin Bradford Hill, and others that smoking causes lung cancer.[3] Like *Science—The Endless Frontier*, *Smoking and Health* has just two layers of organization.

REGINALD MOUNT DESIGNED this antismoking poster for the UK government in the 1960s.

In writing the report, the RCP aimed to influence UK government health policy and to help physicians advise their patients on the risks of smoking. But the RCP also wanted to change people's habits. Perhaps for that reason, the report was printed as a book, with book-sized margins, left and right justification, hyphenation of words at the ends of lines, and a book-typical number of characters per line (about 70 including spaces).

People are used to consuming tens of thousands of words in book form. For-profit publishers are skilled at arranging words on pages in ways that don't strain readers' eyes. *Smoking and Health*, the book, is 70 pages long. Within a year of publication, it had sold 33 000 copies in the UK and 55 000 in the US.

Unfortunately, too many book-length reports from nonprofit publishers—notably the National Academy of Sciences—violate the principles of good, reader-friendly design. They inflict on readers page after page of front matter, their lines are eye-strainingly long, and bullets and boldface abound.

If you want people to read your report, make it readable.

References

1. V. Bush, *Science—The Endless Frontier: A Report to the President on a Program for Postwar Scientific Research*, NSF (1945; reprint, 1990).
2. E. K. Gibson, *Science* **369**, 258 (2020).
3. Royal College of Physicians, *Smoking and Health: Summary of a Report of the Royal College of Physicians of London on Smoking in Relation to Cancer of the Lung and Other Diseases*, Pitman Medical (1962).

Reaching negative CO_2 emissions

David Kramer's article "Negative carbon dioxide emissions" (PHYSICS TODAY, January 2020, page 44) provides an excellent overview of the pros and cons of several climate-ameliorating interventions in the global carbon cycle. But it overlooks what ought to be at the top of our list: protection of natural carbon sinks.

Over the past decade, natural sinks have removed from the atmosphere about 5 gigatons of carbon per year, with approximately three-fifths going to the oceans and the rest to terrestrial ecosystems. That removal rate is about one-half of annual anthropogenic emissions worldwide. And nature does it for free!

That natural sink strength is far greater and far cheaper than any engineered scheme can promise to deliver over the coming two or three decades. If the strength is maintained, a 50% reduction in today's emissions would stabilize atmospheric carbon dioxide; with a further reduction in emissions, atmospheric levels would decline over time, albeit at an ever-decreasing rate.

Unfortunately, natural sinks are threatened today by a combination of deforestation, soil erosion, ocean acidification, agricultural malpractices on prime land, increasing exploitation of poorer quality lands for food production, and climate change itself. Yet on farmland, protection and augmentation of natural sinks can even increase crop yields.

The most important thing we can do to sequester carbon is to prevent the degradation of existing, priceless, and cost-free natural sinks. Combined with rapid deployment of renewable energy, they might give us a chance to prevent climate catastrophe.

John Harte
(jharte@berkeley.edu)
University of California, Berkeley

I appreciate David Kramer's informative article "Negative carbon dioxide emissions" (PHYSICS TODAY, January 2020, page 44). It appears that negative emission technology (NET) will be needed. However, I am puzzled by one phrase—"achieving Paris goals without retarding economic growth." Isn't it obvious that perpetual economic growth in our finite terrarium/aquarium is not possible? Long-time PHYSICS TODAY readers will remember several items by Albert Bartlett about exponential growth (see, for example, PHYSICS TODAY, July 2004, page 53, and March 1994, page 92). Such growth—economic and other—is a primary driver of increasing carbon dioxide emissions and thus of climate change.

Kramer quotes Julio Friedmann: "We have to create an industry the size of the oil and gas industry that runs in reverse." The oil and gas industry generates its output to make a profit. The "reverse output" of the NET industry—tens of gigatons of CO_2 sequestered annually—is not a marketable product to be sold for up to $100 per ton. A carbon fee may be needed to fund NET.


ISTOCK.COM/FLORENT ROLS

Would use of CO_2 for enhanced recovery of oil make up any significant fraction of tens of gigatons per year? How much CO_2 would be released in burning the additional oil produced? Does the CO_2 come back up with the oil? Then what?

Kramer mentions "improved forest management" but doesn't explain it. In the western US, politicians and others often use that phrase as code for "more logging." (They usually avoid the word "logging" in favor of "thinning," "forest health," or similar language.) A purported goal is to reduce the fuel available to forest fires. Proposals to increase CO_2 storage by reforestation and large-scale tree planting on public land are incompatible with demands for forest-fire fuel reduction. Studies by wildfire scientists and ecologists[1,2] show that such fuel reduction is generally ineffective in reducing large fires and is ecologically damaging. It is not improved management.

References

1. D. A. DellaSala, testimony before the US House Natural Resources Committee, Subcommittee on Oversight and Investigations, 27 September 2017, p. 3.
2. T. Schoennagel et al., *Proc. Natl. Acad. Sci. USA* **114**, 4582 (2017).

Dick Walton
(dwalton@centurylink.net)
Billings, Montana

Carbon pricing needs a dividend

Media outlets have emphasized how our changing climate fuels such tragic events as the blazing Australian bushfires in 2019–20. One should also remember that policies can be enacted to address the daunting challenge of climate change. David Kramer's story "Should carbon emissions be taxed or capped and traded?" (PHYSICS TODAY, December 2019, page 28) provides a timely comparison of market-based policies aimed at reducing greenhouse gas emissions. One approach that is catching unprecedented attention among economists and members of the US Congress is a carbon fee coupled with a dividend.

I think about the economics of carbon pricing in terms of uncertainty. In physics, we are used to thinking of trade-offs in uncertainty when measuring conjugate variables such as position and momentum. Economists inevitably trade off uncertainty in price and quantity when designing carbon-pricing policies; a carbon fee addresses the former, and cap and trade the latter. The distinction may sound pedantic, but it matters when courting stakeholders.

Ensuring certainty in carbon's price is a practical political move. Although we must reduce our carbon emissions, the exact quantity of that reduction can afford some uncertainty. The benefit of fixing carbon's price is the ability to minimize economic risk,[1] a choice that businesses generally prefer. Under the carbon fee policy, a fixed fee is charged for each ton of emitted carbon dioxide and is gradually increased each year.

Kramer's story briefly mentions what might happen to the revenue generated from the carbon fee. Should it be used at the government's discretion or returned as an equal dividend to each citizen?

Incidents of resistance to carbon pricing point to the need for a dividend. Kramer cites the French yellow vest protests as an example of how a faulty policy can cause social unrest. But the crucial lesson is that any carbon-pricing policy must adequately address economic inequality. By returning a dividend equally to all taxpayers, the policy becomes overall progressive and effectively revenue neutral—a more politically viable option in line with a conservative desire for limited government.

Economists are in nearly universal agreement that carbon pricing is best accompanied by a dividend. A historic statement published in the *Wall Street Journal*[2] made the consensus explicit. Among the signatories are 27 Nobel laureates, 4 former chairs of the Federal Reserve, and more than 3500 US economists.

Beyond that remarkable level of unity among economists, more than half a dozen carbon-pricing bills were introduced in Congress last year. The one with the most congressional cosponsors, by about an order of magnitude, is the Energy Innovation and Carbon Dividend Act of 2019 (HR 763). Last October Columbia University's Center on Global Energy Policy released an assessment of HR 763 that projects emissions reductions by 2030 will "exceed the US commitments to the Paris Agreement" and that the "substantial revenue" returned as a dividend will generally benefit low- and middle-income households more than the fee hurts them.[3]

Despite growing support for the bill, I sympathize with Kramer's concern regarding "today's polarized US political climate" and the likelihood of opposition to carbon pricing. But I have found that participating in our democracy is an antidote to that pessimistic outlook. Passing national carbon-pricing legislation requires a sufficient level of political will, which we create by asking our elected representatives to support HR 763.

References

1. C. Hepburn, *Oxf. Rev. Econ. Policy* **22**, 226 (2006).
2. "Economists' statement on carbon dividends," *Wall Street Journal*, 16 January 2019.
3. N. Kaufman, J. Larsen, P. Marsters, H. Kolus, S. Mohan, *An Assessment of the Energy Innovation and Carbon Dividend Act*, Columbia University (2019).

Adam Reed
(adam.p.reed@protonmail.com)
Citizens' Climate Lobby
Longmont, Colorado

Time, the revelator

I am now retired, but throughout my career as a professor I consistently argued that student opinion forms collected in the last week of a course were nearly useless for the evaluation or improvement of teaching effectiveness, though the free-form comments were occasionally useful or at least amusing. I was surprised that an ideal mechanism to evaluate teaching—which I've championed for at least 25 years—was not included in Toni Feder's article "Reevaluating teacher evaluations in higher education" (PHYSICS TODAY, January 2020, page 24).

Technology has made it simple to keep track of students who took a particular class and to send them an email questionnaire about it a few years later. Did that course have a positive effect on their education—for example, on their preparedness for subsequent courses—or on their careers? Answers to questions like the following would help make that determination:

- Did you find the material taught in Physics 000 by Professor X useful in subsequent courses?
- Do you ever go back to your notes or textbook from that class to review information because it is important in another course or in your current job?
- Did the solution to a problem or exercise in Professor X's course turn out to be applicable in your current position?
- Have you applied a technique learned in Professor X's class to solve a different problem you've subsequently encountered?
- In hindsight, would you recommend that a student take Physics 000 with Professor X or with some other instructor?

Students would be unable to answer those questions in the last week of class, but years later they could make a much better and more relevant assessment. In addition, such evaluations are not needed in the short term. By the time a faculty member is coming up for tenure in their seventh year, many students will have taken their courses and graduated.

I made a similar suggestion when I was on my college's sabbatical committee. A faculty member is generally allowed to take a sabbatical after receiving tenure at seven years of service. When eligible for the next sabbatical, that faculty member should be asked what benefits from the last one were realized during the intervening years. That suggestion was received with the same lack of enthusiasm as my suggestion for evaluations of teaching effectiveness that used a longer-term student perspective.

Although a student will generally know if a professor is unprepared for class or is irresponsible in grading and returning assignments, that kind of information would probably be brought to a department chair's attention during the semester. A common saying best summarizes what is more important for faculty evaluations: "You don't get what you expect, you get what you inspect."[1]

If instructors know that their performance will be judged by the impact their course has on students' futures, the debate will change from a survey that assesses classroom experience to a focus on the true purpose of that experience.

Reference

1. See, for example, K. R. Smith, *Energy Sustain. Environ.* **11**(2), 3 (2007).

Steven Garrett
(sxg185@psu.edu)
Pennsylvania State University
University Park

Units, for good measure

In his editorial in the March 2020 issue of PHYSICS TODAY (page 8), Charles Day commented that Lord Rayleigh cited cubic millimeters for the volume of pipes but inches for their length. In the US, we measure gasoline consumption in miles per gallon, whereas in Europe, it is measured in liters per 100 kilometers, the inverse. The European metric unit has dimensions of area, but the US unit is 1/area. What is the meaning of that area? My hybrid car has a consumption of about 40 miles per gallon. The corresponding area—much easier in metric—is about 0.6 mm². That is the cross-sectional area of a gasoline-filled pipe that, if laid along the road, will keep my car going if I scoop up the gas. It may be 2 mm² for a gas-guzzler. I propose measuring fuel consumption in square millimeters.

Michael Albrow
(albrow@fnal.gov)
Fermilab
Batavia, Illinois

Corrections

March 2020, page 45—The equation should read as follows:

$$\underbrace{\frac{1}{v_0}\frac{\partial \eta}{\partial t} + \frac{\partial \eta}{\partial x}}_{\text{Transport}} + \underbrace{\frac{h^2}{6}\frac{\partial^3 \eta}{\partial x^3}}_{\text{Dispersion}} + \underbrace{\frac{3}{2h}\eta\frac{\partial \eta}{\partial x}}_{\text{Interaction with seabed}} = 0,$$

March 2020, page 46—The $L_{ij}^{(k)}$ and $L_{xy}^{(z)}$ terms should have been $\mathcal{L}_{ij}^{(k)}$ and $\mathcal{L}_{xy}^{(z)}$, respectively. PT

CONTACT PHYSICS TODAY

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

Cold-atom lattice bends topological rules

In a periodically driven system, exotic phases can form that have no static counterparts.

The complex quantum interactions of electrons in solids give rise to a plethora of phenomena—most famously, high-temperature superconductivity—that challenge the ingenuity of theorists. Fortunately, experimenters have a path to gaining new understanding of those condensed-matter systems by engineering the same physics in an ultracold atomic gas. Atoms in a trap are easier to tune, control, and measure than electrons in a solid, and cold-atom researchers have built up a toolbox of techniques for mimicking electronic systems. Artificial magnetic fields, for example, can act on neutral atoms the same way real fields act on electrons (see the article by Victor Galitski, Gediminas Juzeliūnas, and Ian Spielman, PHYSICS TODAY, January 2019, page 38).

Now Monika Aidelsburger of Ludwig-Maximilians University Munich, her students Karen Wintersperger and Christoph Braun, and their colleagues are using those same tools to go beyond simulating condensed matter and into a new regime with no material counterpart.[1]

Their cold-atom experiment lies in the realm of topological physics, in which unusual behaviors at a system's edges relate to and are determined by properties of the bulk. (See the article by Arthur Ramirez and Brian Skinner on page 30 of this issue.) In particular, they're studying a class of systems with chiral edge modes: quantum states that carry particles around the system's perimeter, but only in one direction. Ordinarily, edge modes can be quantified by topological invariants called Chern numbers, which can be computed from measurements in the bulk. But the system Aidelsburger and colleagues have realized includes the so-called anomalous Floquet regime, in which edge modes are present but all the Chern numbers are zero.

Experiments on the anomalous Floquet regime should make it possible for experimenters to explore topological phenomena that are impossible to study in equilibrium—not because their manifestations in static solid materials are too complicated, but because they don't exist.

Driving test

The anomalous Floquet regime hinges on the unusual physics of periodically time-varying Hamiltonians—the same kind used to create time crystals (see the article by Norman Yao and Chetan Nayak, PHYSICS TODAY, September 2018, page 40). By convention, periodically driven systems are analyzed stroboscopically: Time is considered not as a continuous variable but in only discrete multiples of the driving period. That technique is blind to the so-called micromotion that occurs over the course of each driving cycle.

But the micromotion can lead to some interesting behavior. For example, consider the driving protocol from Aidelsburger and colleagues' experiment, shown in figure 1. The two-dimensional honeycomb lattice is simply the interference pattern of three lasers. By oscillating the lasers' intensities, the researchers can lower the energy barriers in each direction in turn to give atoms in the lattice a chance to hop from site to site along the red lines. In the conceptually simple limit—when the hopping probability approaches 1 along the red bonds and is zero along the gray ones—an atom in the bulk of the lattice travels counterclockwise around a single hexagon. After two driving periods, the atom is back to where it started, so in that stroboscopic view, the system is trivial: Particles localized in the bulk just stay where they are.

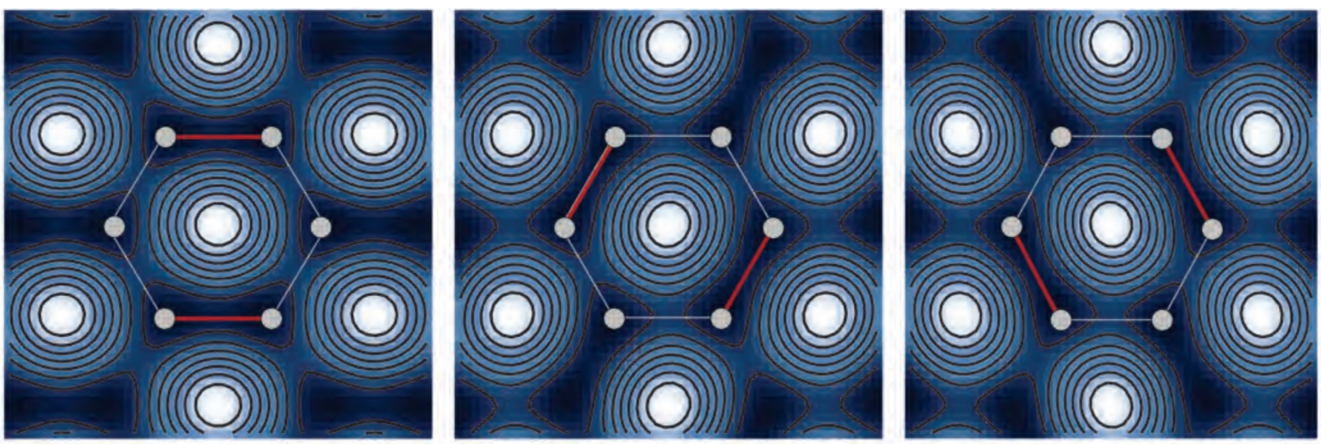


FIGURE 1. TOPOLOGICAL PHYSICS in a honeycomb optical lattice, created from three interfering lasers angled 120° from one another. Modulating the lasers' intensities periodically lowers the energy barriers along each set of parallel edges in turn, as shown by the red line segments. That protocol engenders qualitatively different behavior in the lattice bulk and at its edges. (Adapted from ref. 1.)

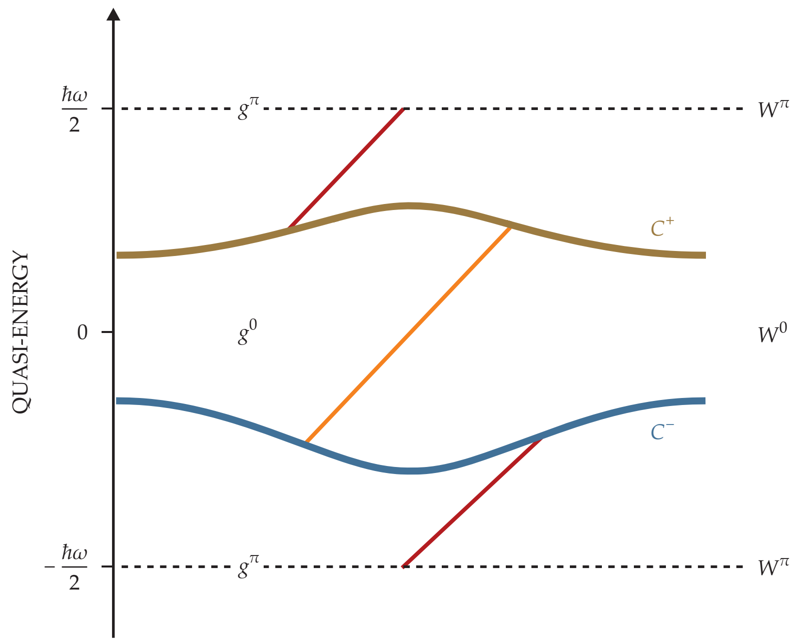


FIGURE 2. ABSOLUTE ENERGY DOESN'T EXIST in periodically driven systems; states have only a quasi-energy that's defined modulo $\hbar\omega$, where ω is the driving frequency. Two energy bands (blue and gold) can be connected by an edge mode across either the zero-energy gap g^0 (orange) or the quasi-energy gap g^π (red). The bands' Chern numbers C^+ and C^- are equal to the number of modes connecting to the band from above minus the number connecting from below. When both modes are present, C^+ and C^- are zero; they no longer define the system's topological state. Instead, the system is characterized by the winding numbers W^0 and W^π, equal to the number of modes traversing each gap. (Adapted from ref. 1.)

Particles that start at certain edge sites, however, don't just stay where they are. If an atom can't complete a trip around its hexagon without running into the lattice edge, it's instead carried clockwise along the lattice perimeter. Despite the system's trivial stroboscopic Hamiltonian, the intracycle micromotion gives it a chiral edge mode. Importantly, though, the researchers' interest in the system isn't limited to that simple example. The system's behavior depends on both the amplitude and frequency of the laser modulations, and its topological properties vary across that 2D phase diagram.

Figure 2 shows a schematic of the system's energetics. Two energy bands, analogous to a valence band and a conduction band, are shown in blue and gold, and edge modes correspond to connections between bands. Each band has its own Chern number, equal to the number of connections it makes to higher-energy bands minus the number it makes to lower-energy bands.

Ordinarily, if a system has any edge modes, the Chern numbers must be nonzero for at least the lowest- and highest-energy bands that participate in the modes. Furthermore, if one knows all the Chern numbers, one can deduce the system's entire edge-mode structure. But in a periodically driven system, that reasoning breaks down.

Because driven systems are perpetually out of equilibrium, and because energy can be pumped in or out during the driving cycle, their eigenstates no longer have absolute energies, only quasi-energies defined up to an integer multiple of $\hbar\omega$, where ω is the driving frequency. The quasi-energy spectrum is periodic, so the bands can connect in two ways, as shown by the red and orange lines in the figure. When both modes are present, both bands' Chern numbers are zero, just as if both modes were absent.

Winding roads

The theory of anomalous Floquet systems was first described[2] in 2010, and by

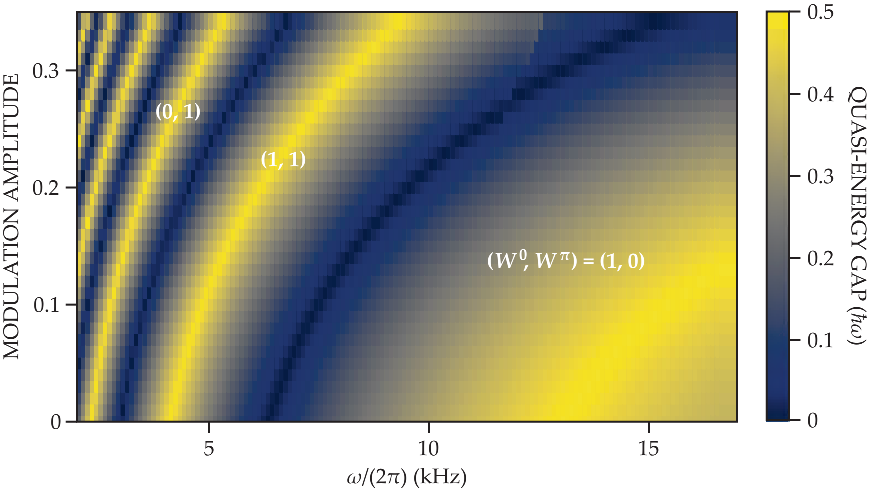


FIGURE 3. MAPPING THE PHASE DIAGRAM of the driven atomic system. Interferometric measurements probe the smaller of the two energy gaps—g^0 and g^π in figure 2—as a function of the modulation amplitude and frequency ω. When one of the gaps is zero, the bands touch and an edge mode can appear or disappear. The dark blue lines therefore mark the boundaries between topological phases. The anomalous Floquet regime, in which both winding numbers W^0 and W^π equal 1, is the second yellow region from the right. (Adapted from ref. 1.)

2017 the anomalous regime had been realized experimentally in arrays of photonic waveguides.[3] But only now, with Aidelsburger and colleagues' cold-atom experiment, is it possible to readily tune and characterize such a system. Because the topology is no longer fully defined by the Chern numbers, one must turn to a different set of topological invariants, the winding numbers W^0 and W^π, which are equal to the number of edge modes traversing each gap. And Aidelsburger and colleagues have figured out a way to measure them.

They start with a technique called Stückelberg interferometry: They prepare an atomic wavepacket that lies partly in each energy band, let it propagate for a time, and probe how it interferes with itself. That measurement gives the smaller of the quasi-energy gaps g^0 and g^π, and when mapped for all values of the modulation amplitude and frequency, the measurements yield the phase diagram shown in figure 3.

Edge modes can appear or disappear—and Chern and winding numbers can change—only when an energy gap is zero and the bands touch. The dark blue lines in the figure therefore represent boundaries between topological phases with different numbers of edge modes. The interferometry measurements by themselves, though, don't indicate which of the energy gaps closed or whether the associated winding number increased or decreased.

To fill in that information, the researchers first noted that the high-frequency limit is equivalent to a static system: The quasi-energy spectrum is wide, the gap g^π is large, and the winding number W^π must equal zero. As the frequency is decreased, the gap g^π is the first to close, then g^0, and so on. Finally, by measuring the Hall deflection on either side of each phase transition (see the article by Joseph Avron, Daniel Osadchy, and Ruedi Seiler, PHYSICS TODAY, August 2003, page 38), they deduced the transition's topological charge—that is, whether crossing the phase boundary causes the number of edge modes to increase or decrease by one. By tracking those changes across the phase diagram, they deduced both winding numbers for each phase, and they demonstrated that the anomalous Floquet regime is the second phase from the right.

Into the unknown

Everything the experimenters have found was already anticipated by theory. But that's because their experiment focused on the system in its simplest, most theoretically tractable form: noninteracting atoms in a perfectly uniform lattice. Adding more complicating factors, such as lattice defects or interactions, quickly takes the system beyond theorists' abilities to reliably model it and into uncharted experimental territory.

Especially interesting is the effect of disorder. One of the hallmarks of topological properties is that they don't change when the system is slightly deformed, and theory tentatively predicts that the anomalous Floquet phase, because its Chern numbers are all zero, should be especially robust. By projecting an optical speckle pattern onto their lattice to alter the trap depths and barrier heights, Aidelsburger and colleagues will investigate just how much disorder they can introduce without disrupting the phase's topology.

The experiments so far have focused on measuring the band structure of the periodically driven lattice, irrespective of how those bands are filled by atoms. Another important goal, then, is to figure out how to completely fill one of the system's energy bands—a different and more difficult task than completely filling the lattice—and leave the other band empty.[4] If the researchers can stabilize the system in such a state, they will have realized a new phase of matter, an anomalous Floquet insulator.

Strikingly, Aidelsburger and colleagues managed to completely characterize their system's edge modes without ever looking at the edge modes themselves. In fact, they can't study the edge modes directly because the system lacks well-defined edges—like most optical lattices, the optical intensity fades away gradually in all directions. So one more item on their to-do list is to engineer a new optical potential with a sharp interface between the lattice and the vacuum. With it, they hope to get a reassuring glimpse of chiral edge transport in action.

Johanna Miller

References

1. K. Wintersperger et al., *Nat. Phys.* (2020), doi:10.1038/s41567-020-0949-y.
2. T. Kitagawa et al., *Phys. Rev. B* **82**, 235114 (2010).
3. S. Mukherjee et al., *Nat. Commun.* **8**, 13918 (2017); L. J. Maczewsky et al., *Nat. Commun.* **8**, 13756 (2017).
4. M. S. Rudner, N. H. Lindner, https://arxiv.org/abs/1909.02008.

An Alaskan volcano, climate change, and the history of ancient Rome

In the first century BCE, political instability set the Roman Republic on edge. Researchers have found evidence of an eruption that may have catalyzed the republic's end.

FIGURE 1. THE 10-KM-WIDE CALDERA on Alaska's Umnak Island was formed during the 43 BCE eruption of the Okmok volcano. The massive eruption produced some of the most extreme Northern Hemisphere weather conditions of the past 2500 years. (Courtesy of Kerry Key, Columbia University.)

The assassination of Julius Caesar on the Ides of March in 44 BCE marked the beginning of a 17-year power struggle for control over the greater Mediterranean region. Roman law remained in force, but societies were in tumult. Civil wars broke out, the Senate and its consuls were ousted, and in 31 BCE Julius's grandnephew and heir Octavian emerged as the first emperor of the Roman Empire.

Accounts of the time reveal that inclement weather, crop failure, widespread famine, and disease accompanied the political upheaval. Recent climate reconstructions have borne out those accounts: On average, 43 and 42 BCE were among the coldest of the past 2500 years in the Northern Hemisphere.[1]

A team of climate scientists, volcanologists, and historians led by Joseph McConnell of the Desert Research Institute in Reno, Nevada, has now compiled persuasive evidence pinpointing the source responsible for the extreme environmental conditions: a powerful eruption from Alaska's Okmok volcano (figure 1), 9300 kilometers from Rome.[2]

To reach that conclusion, the researchers measured the volcano's ancient fallout sequestered in six well-dated Arctic ice cores. Their geochemical analysis of volcanic tephra—small shards of cooled magma—in the cores revealed a near-perfect match to basalt found around the volcano today. Climate proxies, such as tree rings and stalagmite layers, provided a complementary record of temperatures and rainfall at the time of the ancient event. So did historical chronicles and climate modeling.

The research represents one of a growing number of interdisciplinary projects that seek to reconstruct an era's climate and better understand the history of a civilization in its proper environmental context.

Fire and ice

Several factors influence the variability of Earth's climate. The Sun's activity rises and falls with its 11-year sunspot cycle. Earth's tilt, spin, and orbit likewise oscillate through so-called Milankovitch cycles, which influence the amount and distribution of solar energy that reaches Earth's surface (see the article by Mark Maslin, Physics Today, May 2020, page 48). The Milankovitch cycles induce ice ages, which take thousands of years to develop—hardly a rapid response. By contrast, volcanic eruptions are sudden, dramatic events that can shock a society for years.

During a large eruption, billions of tons of rock, gas, and ash are thrust into the atmosphere. Small, insoluble particles of tephra rain out within days to weeks. But the gas, much of which is sulfur dioxide, can reach the stratosphere, where it may linger for years. In the thin dry air, SO_2 slowly becomes oxidized to sulfate aerosols. Depending on the eruption's location, those sulfates can get caught up in stratospheric winds, which disperse them throughout one or both of Earth's hemispheres. Extremely reflective, the sulfates scatter incoming solar radiation and cool Earth's surface. But eventually they, too, fall out, some settling on snow near one of the poles.

As the snow becomes compressed into ice over the millennia, gases and molecular compounds become trapped, layer by layer, and form a rich, time-stamped record of the atmosphere's chemical composition. Industrial pollution, dust, pollen, ash from biomass burning, and volcanic residue can all be found there. Since the 1960s, researchers have drilled cylindrical ice cores in glaciers and ice sheets to mine that archive.

Early on, the ice cores were cut into short slices, and individual layers were melted and analyzed one by one as a function of ice depth. But the technology has matured; today, researchers can continuously melt and analyze ice cores simultaneously for much more accurate dating. "Each year is much more clearly resolved," McConnell says. "And because nearly everything we measure in the ice shows seasonal variation, we can use multiple parameters to identify those annual layers." His system, shown in figure 2, holds an ice core over a heated ceramic plate. Grooves in the plate channel meltwater from the innermost part of the core to a mass spectrometer to measure the concentration of dozens of elements.

Lead pollution and volcano hunting

McConnell and his collaborators did not set out to look for the source of sudden

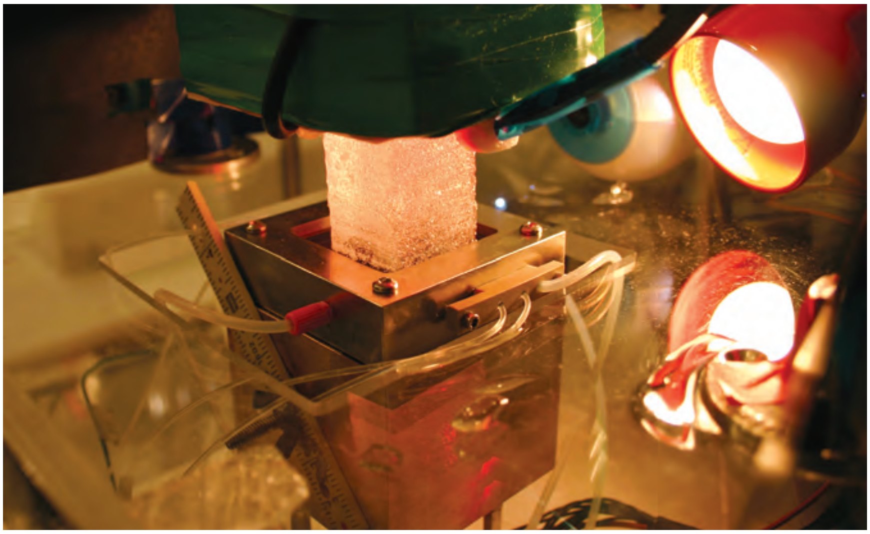


FIGURE 2. AN ICE-MELTER PLATE holds a 3 cm × 3 cm block from an ice core. As the ice melts, its water is pumped into two mass spectrometers and vaporized for elemental and chemical analysis in real time. The system can resolve concentrations at the level of parts per quadrillion and age as a function of depth to within ±2 years. (Courtesy of Joseph R. McConnell.)

climate change in the ancient Mediterranean region. Two years ago, they were analyzing lead pollution in ice cores for a project on the history of silver production. Silver mining and smelting operations waxed and waned in the ancient world. (About 10 000 g of argentiferous lead ore was used to produce 1 g of silver.) McConnell's team used the rise and fall of lead's concentration in the cores as a marker of ancient economic activity, war, and plague.[3]

While reviewing data for that project, McConnell and his former postdoc Michael Sigl (now at the University of Bern) came across an unusually well-preserved layer of tephra in one of the ice cores used in the pollution study. Excited about finding possibly unknown volcanic eruptions, they decided to investigate it further. Their subsequent mass spectrometry measurements of ice cores (figure 3) revealed volcanic fallout from two distinct eruptions—the first starting in early 45 BCE and the second in early 43 BCE, almost a year after Julius Caesar's assassination. The first eruption probably took place in Iceland, the collaboration surmised. But it was short-lived and likely had little climate impact. The second, attributable to the Okmok volcano, was a whopper: It had a greater fallout of sulfuric acid, which persisted for 2.5 years.

The sharp spike in 3–10 µm insoluble tephra shards that McConnell and Sigl had spotted early on coincided with the early stages of the 43 BCE sulfur peak. The temporal narrowness of the spike pinned the eruption to January or February of that year, before the growing season in the Northern Hemisphere.

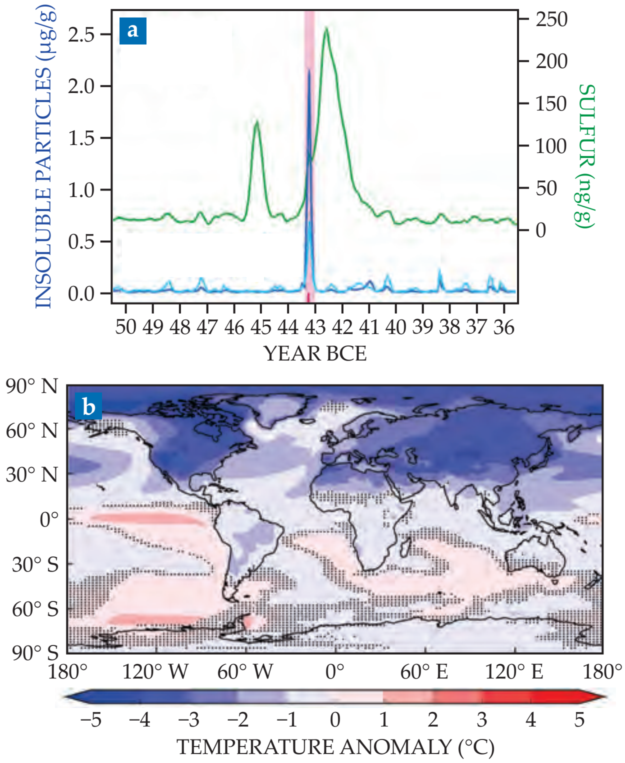


FIGURE 3. ICE-CORE AND CLIMATE ANALYSES. **(a)** Continuous mass spectrometry of sulfur and insoluble-particle (tephra) concentrations—in units of nanograms and micrograms of material per gram of water, respectively—synced to the years that volcanic fallout settled into snow. **(b)** This snapshot of an air-temperature simulation captures the average air-temperature anomalies during 43 and 42 BCE. As inputs, the researchers relied on the Okmok eruption's location, timing, and sulfur yield to estimate its cooling effect and the extent and persistence of the climate response into the late 30s BCE. (Adapted from ref. 2.)

Coauthor Gill Plunkett of Queen's University Belfast conducted electron microprobe analyses on 35 tephra shards filtered from the melted ice cores to confirm the volcano's identity.

Earth has about 1500 potentially active volcanos, not counting the belt of spreading centers on the ocean floor. But their magmas have geochemical fingerprints that reflect each one's distinct location in Earth's crust. The geochemistry of the ice-laden shards and Okmok's contemporary basalt matched well enough for Plunkett to rule out several other possible volcanos.

The 2.5-year spread of the volcanic fallout coincides with air-temperature anomalies recorded in tree rings and cave stalagmites. Those climate proxies confirmed that the eruption's effects were pervasive in the Northern Hemisphere. The ring widths of temperature-sensitive trees in Scandinavia and Austria each revealed summertime cooling of more than 2 °C in 43 and 42 BCE. Radiocarbon dating of stalagmite layers in China's Shihua Cave showed a similarly pronounced temperature drop. And a rare frost ring found in bristlecone pine trees from the mountains of North America confirmed below-freezing temperatures there in early September of 43 BCE.

The researchers' simulations suggest that in some regions those cooling trends persisted into the early 30s BCE. Indeed, the simulated temperature dropped by as much as 7 °C in parts of southern Europe and northern Africa. Rainfall patterns shifted as well, with summer precipitation levels some 50–120% above normal throughout southern Europe and autumn precipitation in some regions reaching 400% above normal.

The Nile fails to flood

The strong temperature gradient between the Northern and Southern Hemispheres could have shifted the intertropical convergence zone, the region where tropical trade winds meet. That shift, in turn, would have moved the East African monsoon rainfalls southward. The headwaters of the Blue Nile in the Ethiopian Highlands are the source of more than 85% of the summer floods that year after year delivered irrigation and silt to the lower reaches of the Nile in Egypt. With those highlands drier than normal, according to the simulations, little flooding likely occurred in 43 or 42 BCE.

Greek historian Appian recorded the lack of flooding; he cited Cleopatra's reluctance to provide help to Rome due to Egypt's famine and pestilence. Although establishing direct causal linkages between crop failures and political decisions is difficult, McConnell and collaborators argue that Okmok's severe environmental impact likely contributed to social unrest. Rome's interest in Egypt as a breadbasket was magnified, no doubt, by the stress on a largely agrarian society, and Egypt's ability to withstand Rome was diminished by famine.[2] Egypt was absorbed into the Roman empire after Cleopatra's suicide in 30 BCE.

No one knows whether the birth of the Roman Empire would have happened without Okmok's influence. But the extreme climate the volcano produced could have hastened it.

Mark Wilson

References

1. M. Sigl et al., *Nature* **523**, 543 (2015).
2. J. R. McConnell et al., *Proc. Natl. Acad. Sci. USA* **117**, 15443 (2020).
3. J. R. McConnell et al., *Proc. Natl. Acad. Sci. USA* **116**, 14910 (2019).

Mechanically stressed phytoplankton light up

The bioluminescent intensity depends on the rate and amplitude of the microorganism's deformation.

FIGURE 1. SWARMS OF THE DINOFLAGELLATE *Noctiluca scintillans* emit blue light that illuminates the coastline in Taiwan. In historical accounts of nautical adventures, mariners marveled at the milky seas caused by bioluminescence. (Photo by RobbinChang.)

English naturalist Henry Baker in 1753 identified the source of the phosphorescent "burning of the seas," like the glow off the Taiwanese coast in figure 1, that had mesmerized nautical explorers for centuries. His prowess in microscopy led him to identify a unicellular "illuminous animal," *Noctiluca scintillans*, that emits light when disturbed, perhaps by a ship's bow or a rower's oar. Understanding of the biochemical processes that generate luminescence in *N. scintillans* and other bacteria or dinoflagellates in marine and freshwater environments progressed rapidly. But what mechanisms ultimately trigger bioluminescence? American zoologist E. Newton Harvey in 1920 commented, "That problem must be left to the physicist."

The complex series of chemical reactions through which bioluminescent marine organisms produce light begins when calcium ions enter the cell through channels in the cell wall. Biologists hypothesize that fluid flow outside the cell causes the cell's outer wall to stretch, which opens those ion channels. One puzzle, however, arises from experiments that use different methods of stimulating the cell to elicit a bioluminescent flash: How exactly do those methods, which involve both fluid and mechanical perturbation, lead to a distribution of forces over the cell body that cause the cell to stretch?

Maziyar Jalaal, Raymond Goldstein, and colleagues at the University of Cambridge now have taken a step toward solving that puzzle. In a series of experiments, the researchers pinpoint the mechanical stresses that deform the viscoelastic membrane of the dinoflagellate *Pyrocystis lunula* and trigger light emission.[1] Linking macroscopic flows to microscopic cellular processes could improve understanding of the ecology and evolution of bioluminescent organisms.

Poking and prodding

P. lunula has occupied the spotlight for bioluminescence studies for several decades. The large (130 μm long) organism, shown in figure 2, has a rigid cell wall, has limited mobility, has a density greater than that of seawater, and is transparent. Because of those properties, handlers can easily manipulate the dinoflagellate's position and observe its internal activity. Experiments by several research groups in the 1960s demonstrated that stirring or bubbling water around the creature could cause it to emit flashes of light from its central cytoplasmic area.[2] That work motivated quantitative investigations of how fluid motion triggers bioluminescence.

Subsequent studies using improved flow visualization tools have shed light on the process. In the early 2000s, Michael Latz and colleagues at the Scripps Institution of Oceanography estimated the magnitude of stress required to trigger light production in *P. lunula* under different methods of forcing. In one study, they poked a cell at a single point using an atomic force microscope until the cell lit up.[3] In other studies, they pumped water alongside the cell and found that emitted light intensity depends on the shear.[4] And zoophysiology experiments performed by Latz's group have shown how much stress is required to trigger stretch-activated proteins in dinoflagellate bioluminescence signalling pathways.[5]

But there are discrepancies in bioluminescent responses to different forms of stimulation. Those variations are likely based on the unknown distribution of sensors in the cell that respond to mechanical stimulation and on how the sensors' activity is integrated into a whole-cell response.

Jalaal and his colleagues have now performed a quantitative study that could help resolve the discrepancies. Their experiments explored two ways of stimulating *P. lunula* cells to flash. In the first protocol, the researchers pumped a liquid through a microscopic pipette to create a shearing force along the surface of a cell (figure 3a and b). Tiny particles in the liquid traced the flow lines near the cell surface, revealing the velocity and the resulting distribution of forces on the cell. A high-speed camera recorded the

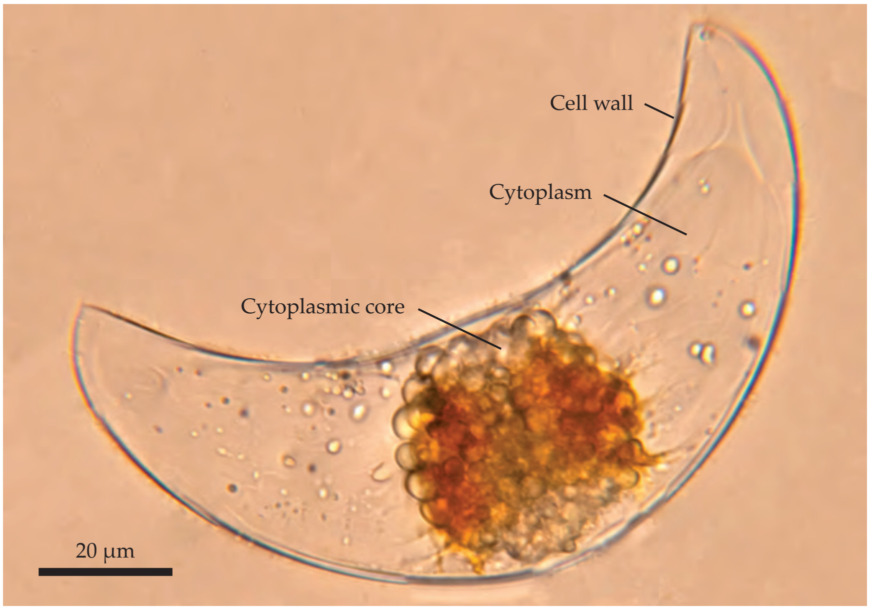
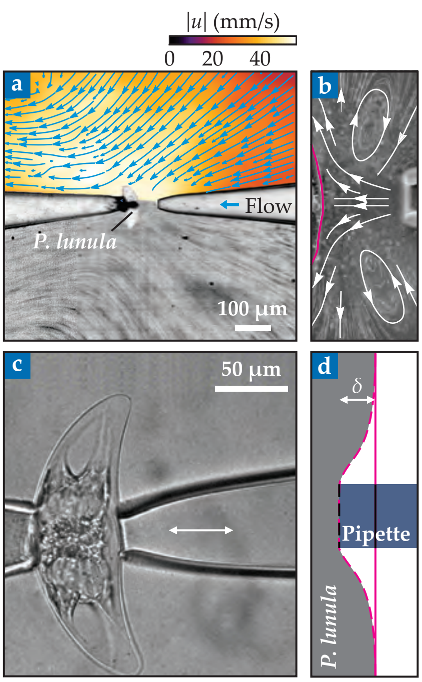


FIGURE 2. A SINGLE CELL LIGHTS UP. The unicellular marine phytoplankton *Pyrocystis lunula* is a common subject for bioluminescence studies. The organism's cytoplasmic core emits flashes of light when external stresses push on the cell wall. (Adapted from ref. 1.)

tracer-particle positions and the cell's distortion and light emissions. Only when the imposed fluid pressure was high enough to sufficiently deform the cell's outer membrane did the cell flash.

In the second protocol, the researchers used two pipettes to hold and gently squeeze the cell, and that squeezing created a highly localized deformation (figure 3c and d). Micromanipulators allowed them to independently vary both the magnitude and rate of the deformation. As with the fluid flow experiment, a high-speed camera recorded the cell's distortion and light emissions.



Sensitive stretches

For both experimental protocols, the imposed forces were increased at a constant rate to achieve a final deformation, after which the force was held fixed. *P. lunula* emitted a quick flash of light that slowly decayed. Observations of light intensity at different deformation rates (10–900 μm/s) and magnitudes (1–10 μm) showed that light intensity depends not only on how much the cell is deformed but also on the speed of the deformation.

FIGURE 3. TWO DIFFERENT MECHANISMS trigger a *Pyrocystis lunula* specimen to produce light. In both experiments, the unicellular phytoplankton sits on the tip of a micropipette. **(a)** Under fluid stimulation, a jet of liquid directed toward the cell at velocity u creates normal and shear forces. The upper half sketches the flow lines (blue), and the color map indicates flow speed u; the lower half shows streaks of particle tracers in the fluid. **(b)** Streamlines map out the flows that ultimately deform the cell wall (red) and elicit light production. **(c)** Under direct mechanical stimulation, a second pipette squeezes the cell. **(d)** The pipette deforms the cell wall (red) by an amount δ until light production ceases. (Adapted from ref. 1.)

By repeating both experiments several times on individual cells, the researchers showed that the maximum light intensity emitted by a cell decreased in subsequent perturbations.

The findings led Jalaal and coworkers to develop a mathematical model that quantifies light production as a viscoelastic response. The model describes the triggering signal in terms of a strain rate that's scaled by the relaxation time of the viscoelastic cell wall. The relationship between that signal and the emitted light intensity depends on the time scale on which the light-emitting biochemical reaction completes and on the time scale on which the cell resets and prepares to flash again. The triggering signal saturates at a maximum value—perhaps when all available ion channels in the cell membrane are open: The signal might be a proxy for the influx of calcium that results from the opening of channels. The model captures the observation that peak light intensity occurs when cells are highly deformed at high speed.

Combined with earlier studies that involved perturbing cells with different mechanical forces or flows, the new observations suggest that a cell may generate the same amount of light in response to a small force spread over a large area or a large force focused on a small area—that is, the light could result from weakly stretching many ion channels in the cell membrane or from strongly stretching just a few. Jalaal notes that further studies will be needed to investigate more realistic environmental fluid-flow conditions and determine "the distribution of forces on the cells and details of fluid-cell membrane interaction."

Bioluminescence is a tool that organisms use to communicate with their surroundings—for example, to scare off predators. An improved understanding of its biochemical and physical origins provides the opportunity to explore its ecological significance across scales and species.

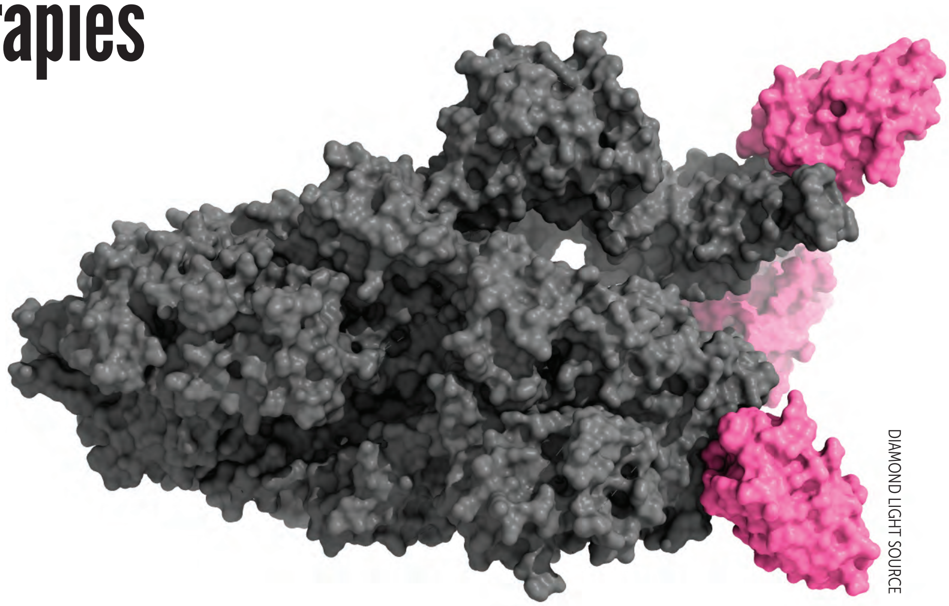
Rachel Berkowitz

References

1. M. Jalaal et al., *Phys. Rev. Lett.* **125**, 028102 (2020).
2. W. H. Biggley et al., *J. Gen. Physiol.* **54**, 96 (1969).
3. B. Tesson, M. I. Latz, *Biophys. J.* **108**, 1341 (2015).
4. P. von Dassow, R. N. Bearon, M. I. Latz, *Limnol. Oceanogr.* **50**, 607 (2005).
5. K. Jin et al., *J. Phycol.* **49**, 733 (2013). PT

Cats and llamas could offer a path to coronavirus therapies

Computation and experimentation also yield possible therapeutic compounds for COVID-19.



THE SPIKE PROTEIN OF THE SARS-COV-2 VIRUS (gray) is shown with three small antibodies (pink) attached to its receptor binding domains. The spike attaches at the left to the viral membrane (not shown).

As the world anxiously awaits development of one or more vaccines to tame the SARS-CoV-2 virus, other research continues at a feverish pace to find effective treatments for the disease it causes, COVID-19. That work, in which physicists and chemists are deeply involved, has made significant strides in the past several months and has turned up a few surprises.

Researchers at the University of Alberta reported at the August virtual meeting of the American Crystallographic Association that a dipeptide-based protease inhibitor used to treat a fatal coronavirus infection in cats also blocks replication of the SARS-CoV-2 virus in samples of monkey lung tissue. Joanne Lemieux, a biochemist at the university, says the antiviral, known as GC373, works by blocking the function of the main protease (M^{pro}), an enzyme that cleaves the polyproteins translated from viral RNA into individual proteins once it enters human cells.

Lemieux says GC373 has been shown to have no toxic effects in cats. Anivive, a California company that develops pet medicines, has applied for US Food and Drug Administration approval to begin trials in humans. Lemieux's group crystallized the M^{pro} in combination with the drug and produced three-dimensional images of how the drug binds strongly to the active pocket on the enzyme. Although GC373 should be effective in its current form, the group is planning further crystallography experiments at the Stanford Synchrotron Radiation Lightsource (SSRL) and the Canadian Light Source to see if a reformulation could optimize it for human use, she says.

At the end of April, SSRL opened a new beamline that has one of the smallest beam sizes and highest brightnesses of any in the world devoted to structural molecular biology and x-ray macromolecular crystallography. "We'll be able to use smaller crystals, collect higher-quality data, get a better signal-to-noise ratio, and collect more data sets per hour" than ever before, says Ian Wilson of Scripps Research, which provided some of the funding for the beamline. Stanford University, several private foundations, and the National Institutes of Health also contributed support.

The search for potential COVID-19 treatments couples experimental and computational efforts. In one example, collaborators at Brookhaven National Laboratory and Stony Brook University developed a complex computational model detailing the sequence of how the viral spike protein attaches to the human host cell. Kerstin Kleese van Dam, director of Brookhaven's Computational Science Initiative, says that when the spike is in its usual down, or inward, position, it is protected from the immune system. The spike becomes more vulnerable when it shifts to the up, or outward, position required to enter the host cell.

The team identified a pocket on the spike that changes shape during the transition from down to up, and compound screening simulations uncovered several molecules that might attach to the pocket and block cell invasion. "It's a bit like wedging a door open," van Dam says. "You might be able to wedge the spike into an upward position, so the immune system can attack it." Chemists are now working to synthesize the predicted molecules. *In vitro* experiments hopefully will confirm the compounds' activity. If the results are positive, light sources or other instruments will probe how the molecules bind to the viral protein.

Jim Brase leads Lawrence Livermore National Laboratory's computational effort on the coronavirus. He says the lab's considerable unclassified computing assets have been running physics-based models to probe the molecular dynamics between proteins and compounds in a search for drug candidates and antibodies that will counter the coronavirus. Antibodies are proteins that bind to and neutralize alien proteins. “You have two very complex molecules and you are trying to see how they come together and what the energy of their binding strength is,” Brase says. “These are typically million-atom calculations, and you have to let [the atoms] jiggle around and settle into their lowest energy states.” The process must be repeated hundreds of times to obtain the average binding energy. “Then you move to the next potential configuration. If you have a thousand of those, you are doing a lot of calculations.”

A commercial vendor manufactures the computer-designed antibodies for Livermore and Sandia National Laboratories, where researchers assay them to determine if they bind to their SARS-CoV-2 targets. Artificial intelligence (AI) is used in winnowing the field of potential new molecules from the nearly infinite number of possible combinations of amino acids.

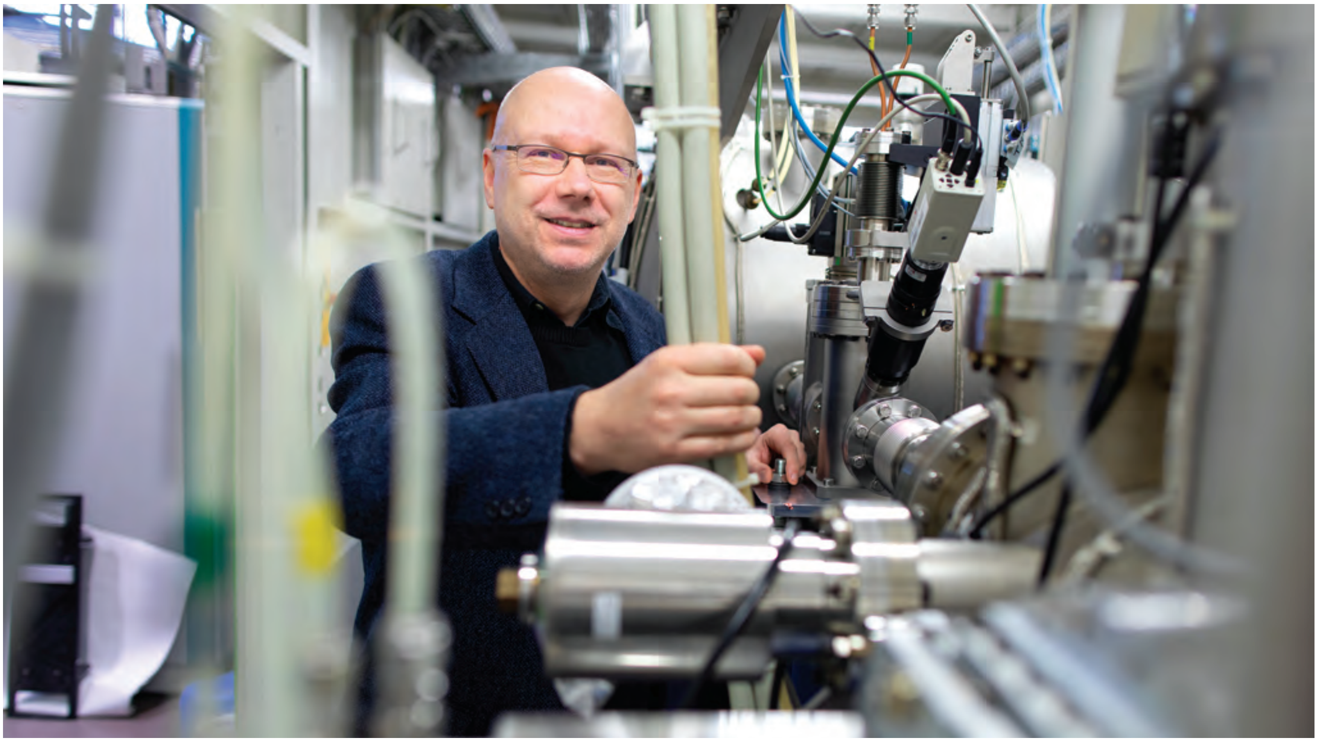
Models developed by the Accelerating Therapeutics for Opportunities in Medicine consortium, comprising Livermore, GlaxoSmithKline, and other university and government entities (see Physics Today, January 2018, page 27), will be used to predict the safety and pharmacokinetics of candidate compounds to become drugs, Brase says.

At Los Alamos National Laboratory, structural biologist Julian Chen is modeling the interactions between the spike protein and the angiotensin-converting enzyme 2 (ACE2) on the human cell surface that is the path to infection. “We know that viruses make use of the path of least resistance to get into the cell,” he says. “SARS-CoV-2 certainly did not evolve specifically to bind to a human ACE2 receptor, so it’s a good enough interaction, but likely not an optimal one.” It should be possible to design molecules that can outcompete the virus in binding with ACE2, thereby preventing infection, he says.

The COVID-19 High Performance Computing Consortium, launched in March, brings together 7 Department of Energy national labs with 8 NSF-funded supercomputing centers, NASA, 11 high-tech industry giants, more than a dozen US universities, and institutions from Japan, South Korea, Sweden, Switzerland, and the UK. Brase says a panel reviews and ranks requests for computational time from investigators around the world. A committee then matches accepted proposals with the consortium members’ available computing resources. Many proposals require more resources than any one member can spare. In those cases, multiple members will contribute time, he says. About 80 projects are underway.

The European Commission is funding an 18-institution coronavirus-fighting consortium known as Exscalate4CoV that includes supercomputing centers at Cineca, Italy’s supercomputing consortium; the Barcelona Supercomputing Center; and Germany’s Jülich Research Center. The Swiss Institute of Bioinformatics contributes 3D models of the viral proteins, and the Fraunhofer Institute for Molecular Biology and Applied Ecology provides a drug-repurposing library of 6000 compounds and biochemical assays. Screenings of crystallized viral proteins in combination with newly discovered molecules are performed at Italy’s

HZB



MANFRED WEISS, director of the macromolecular research group at the BESSY II light source, stands beside a crystallography beamline.

Elettra Synchrotron Trieste, KU Leuven in Belgium, and Poland's International Institute of Molecular and Cell Biology.

FELs enter the picture

Since chloroquine was first proposed as a potential treatment for COVID-19, multiple clinical studies have shown the drug and its derivates to be ineffective and to have harmful side effects. But Claudio Masciovecchio, head of scientific programs at the FERMI free-electron laser (FEL) at Elettra, says he's not ready to write off chloroquine just yet. One recent study showed it to interfere with the spike–ACE2 interaction. He says the FERMI FEL's ability to probe the low-frequency vibrations that are responsible for the biochemical activity of a molecule or drug may settle the issue. Researchers from the University of Bologna, ETH Lausanne, and the University of California, Irvine, are collaborating with FERMI scientists to gain further insight to the interaction. Complementary experiments are to take place this month on Elettra's synchrotron.

SLAC, home to SSRL, is reopening its x-ray FEL (XFEL), the Linac Coherent Light Source (LCLS), this month after an extended shutdown for an upgrade. The LCLS produces x rays in femtosecond pulses at a brightness up to 10 billion times that of a synchrotron. The intensity allows studies of protein crystals that are too small for synchrotrons. The extreme shortness of XFEL pulses allows diffraction patterns to be recorded before radiation damage occurs. Additionally, XFELs can probe crystals at temperatures more relevant to physiological conditions, whereas crystallography at synchrotrons is typically performed at liquid-nitrogen temperatures.

But FELs haven't been able to match the productivity of synchrotrons, where automated sample handling allows high-throughput screening of potential drug compounds that are crystallized together with viral proteins. SLAC scientist Alex Batyuk says the LCLS is now being outfitted to speed experiments. "We know beam time at XFELs is very limited, and we'd like to make it more available for such critical and medically relevant targets, especially at this time," he says.

Progress at light sources

A number of promising compounds for drug treatments have been identified at the UK's Diamond Light Source, which, in addition to screening compound libraries and making the results publicly available, since March has been screening libraries of existing drugs and compounds approved for use in humans for activity against several SARS-CoV-2 proteins. (See PHYSICS TODAY, May 2020, page 22.) Martin Walsh, Diamond deputy director of life sciences, says the results won't be released until the molecules can be confirmed to be safe. "Fingers crossed, we hope to have something in the public domain in the autumn," he says.

In one collaboration with Diamond and Oxford University, the startup Exscientia conducted an initial AI-enabled computational screening of the 15 000 compounds in the ReFRAME drug-repurposing library, which is funded by the Bill and Melinda Gates Foundation. The light source was used to discern the structures of the selected molecules combined with viral proteins.

Diamond also joined forces with the COVID Moonshot, a crowdsourced drug development effort spearheaded by PostEra, a San Francisco–based medicinal chemistry design startup. In that effort, Diamond assessed the activity of more than 900 compounds that were synthesized by Moonshot partners from a list of 10 000 candidates proposed by hundreds of researchers. Diamond obtained structures of 129 of those compounds combined with proteins.

An initial screening of 1200 selected commercially available compounds at the BESSY II light source in Berlin has found "a couple of relatively weakly bound hits," says Manfred Weiss, director of the lab's macromolecular research group. Although those small compounds "are useless for any form of therapy," he says, "they serve as the basis to find larger compounds that bind more strongly and exert some biological effect." Of the 25 to 30 larger follow-up compounds procured, 5 or 6 bind more strongly, he says. Collaborators at the University of Lübeck will next assay them for biological effect.

Llamas' contribution

Finding or manufacturing antibodies may present a more expedient path to a COVID-19 treatment. Antibodies don't enter cells, so they have fewer safety issues, and potential side effects are of less concern than those of drugs. Compared with a drug–protein interaction, an antibody has the entire surface area of the virus to which it could potentially bind, notes Sean McSweeney, director of Brookhaven's Center for Biomolecular Structure.

At Brookhaven's National Synchrotron Light Source II, researchers are performing small-angle x-ray scattering on antibodies in solution. Though the technique produces lower-resolution images than crystallography does, changes to the buffers and other solution components can alter the scattering pattern and yield information about the interactions. "If you have a potential antibody and a target, you can look at them separately through scattering, and together in solution you will see an envelope and a shape that is different," says McSweeney.

Diamond is part of a UK collaboration that last month reported a potential COVID-19 therapy in so-called nanobodies derived from llamas. The surprising contribution of the South American camelid arises from their antibodies' small size: about one-quarter that of humans. That feature increases opportunities for the antibodies to fit into the binding pockets of antigens, in comparison to more complex human antibodies. Nanobodies are more stable, so they could potentially be stored for longer periods after production. And they can be delivered directly to the lungs by an inhaler.

Research dating to the 2002 outbreak of the severe acute respiratory syndrome (SARS) coronavirus found that llamas developed antibodies when exposed to that virus's spike protein. In a 13 July paper in *Nature Structural and Molecular Biology*, scientists at the Rosalind Franklin Institute, Oxford University, Diamond, and Public Health England reported producing a nanobody that bound tightly to the spike protein of SARS-CoV-2 in the laboratory.

Experiments involving camelid antibodies and the spike protein also are underway at the Australian Synchrotron, where the first structure of the SARS-CoV-2 nonstructural protein-9 was produced. Dene Littler and colleagues at Monash University have identified a potential binding site on the NSP9 protein that is thought to be involved in RNA binding.

PAC-MAN

A different approach to COVID-19 therapeutics is being developed at Stanford and Lawrence Berkeley National Laboratory (LBNL). Using the gene-editing tool CRISPR, Stanford bioengineering professor Stanley Qi last year developed a prophylactic antiviral for influenza he called PAC-MAN. In March, Qi began working with Michael Connolly, a principal scientific engineering associate at LBNL's Molecular Foundry, to develop a system for delivering PAC-MAN to the cells of COVID-19 patients.

PAC-MAN is composed of a virus-killing enzyme and a strand of guide RNA, which commands the enzyme to destroy specific nucleotide sequences in viral genomes. Connolly has been developing synthetic molecules known as lipitoids that can deliver PAC-MAN to patient cells by encapsulating the enzyme inside tiny nanoparticles the size of a virus.

In April, Qi and Stanford colleagues demonstrated a type of lipitoid that self-assembled with DNA and RNA into PAC-MAN carriers in a sample of human epithelial lung cells. According to Qi, the system reduced the amount of synthetic SARS-CoV-2 in solution by more than 90%. The Stanford team plans to test in an animal model the PAC-MAN/lipitoid system against a live SARS-CoV-2 virus. The interdisciplinary team will be joined by collaborators at New York University and Karolinska Institute in Stockholm.

Connolly says other efforts at the Molecular Foundry include finding new sensors and viral binding materials that could aid in diagnostics and new materials for personal protection equipment.

For obvious reasons, most coronavirus research has been with recombinant viral proteins instead of the complete virus. While some researchers, including Lemieux and her colleagues at the University of Alberta, were able to test their compounds against the live virus, others, including Masciovecchio, say an inability to work with SARS-CoV-2 has hindered their research. "We are working on protocols which have to be accepted by the health authorities to bring the real virus into Elettra," he says. "It would be really helpful to measure real viruses. It's safe so far as the sample container is secure."

Regardless of their success in fighting the current outbreak, researchers say their work will be highly useful in other viral outbreaks that are sure to come. The main protease of SARS-CoV-2, for example, differs little from that of earlier coronaviruses. "If we find a drug that deactivates the protease of CoV-2, it more than likely deactivates the original SARS as well," says Diamond's Walsh. "And if another similar coronavirus SARS virus emerged, we could quickly have a therapeutic against that."

David Kramer

Particle physicists hash out long-term strategy for Europe

Among the goals are to pick a Higgs factory, carry out R&D on accelerators and detectors, conduct feasibility studies, and improve environmental sustainability.

CERN

THE 100-KM-CIRCUMFERENCE FUTURE CIRCULAR COLLIDER (orange circle) would pass beneath Lake Geneva. It could start out as an electron–positron collider and be upgraded to a 100 TeV hadron collider. A feasibility study on whether to go ahead with it is a top priority in the 2020 European Strategy for Particle Physics. The blue circle shows the position of the Large Hadron Collider.

The European particle-physics community aims to build an electron–positron collider for precision studies with the Higgs boson. And it plans to determine the feasibility of an estimated €21 billion ($24 billion), 100 TeV hadron–hadron collider, known as the Future Circular Collider (FCC), at CERN; as a first stage, the FCC tunnel and infrastructure could be used as a so-called Higgs factory. Those are two top priorities in the latest European Strategy for Particle Physics, which the CERN Council adopted unanimously in June.

In parallel with its pursuit of future colliders, the European particle-physics community should press ahead with established activities, the strategy says. Foremost among them are the high-luminosity upgrade to CERN's Large Hadron Collider (LHC) and support and participation in the neutrino programs in Japan and the US—in particular, the Long-Baseline Neutrino Facility and its associated Deep Underground Neutrino Experiment.

The strategy envisions Europe collaborating with partners around the globe while at the same time ensuring its "continued scientific and technological leadership" in accelerator-based particle physics and related technologies. To realize the strategy goals, and to avoid a large gap in the availability of operational colliders, a quick decision among four possibilities for a Higgs factory is necessary. Additionally, the feasibility studies for the FCC should be completed within seven years to make solid recommendations possible in time for the next strategy update. The current strategy is the second update to the original from 2006.

Thousands of scientists contributed over two years to the strategy. The approval process "was sensitive, and we discussed each comma," says Siegfried Bethke of the Max Planck Institute for Physics in Munich and a member of the CERN Council. That body consists of a scientist and a government delegate from each of the laboratory's 23 member states. Halina Abramowicz of Tel Aviv University chaired the strategy process. "We hope our recommendations will put wind in the sails to move these ideas forward," she says.

A Higgs factory?

"The obvious place to look for deviations from the standard model is a dedicated study of the Higgs boson, because it couples to everything," Abramowicz says. That's why the strategy's first recommendation is a Higgs factory that would involve clean collisions between elementary particles rather than the more complicated mess produced in hadron collisions.

The Higgs was observed in 2012 at the LHC (see PHYSICS TODAY, September 2012, page 12). Although the particle's existence had been predicted, its mass was not pinned down by theorists. The particle-physics community is more in the dark as to what to look for next, with no clear predictions despite many theories. "We have no reliable theoretical guidance," says the University of Liverpool's Max Klein, who coordinates electron–hadron collider developments at CERN (see PHYSICS TODAY, May 2017, page 29). "This calls for a balanced program, and we should perhaps make fewer promises about discoveries," he says. "Increased energy, luminosity, and diversity in collision processes will all be required to explore nature and search for answers to persistent puzzles of particle physics."

"We need a new physics theory that is more complete than the standard model," says Young-Kee Kim of the University of Chicago. She is leading the current Snowmass study, which has until late 2021 to provide input for the US counterpart to the European Strategy for Particle Physics.

Of the four candidates for a Higgs factory, the design that is furthest along is the International Linear Collider (ILC), which would be sited in Japan. The idea has been around for about two decades (see, for example, PHYSICS TODAY, April 2007, page 26). The ILC would create 250 GeV center-of-mass collisions in a 20 km tunnel. The estimated price tag is \$5 billion–\$6 billion, not including labor. The energy could be increased by extending the length or perhaps by upgrading to new accelerator technologies that have steeper gradients. The strategy document says that the European particle-physics community "would wish to collaborate" if the ILC is realized quickly. "It would be a great addition to the community if Japan carried this project forward," says Abramowicz.

Yasuhiro Okada, an executive director of KEK, says that Japan "welcomes the European statement" and that it "can make a difference" for proceeding with the seemingly stalled ILC. The project requires financial and labor commitments not only by Japan but also by Europe, the US, and other partners. A 22 February 2020 statement on the ILC by the International Committee for Future Accelerators could also nudge Japan to get moving on it, Okada notes.

Work started this summer on a final ILC engineering design. In parallel, Okada expects Japan's Ministry of Education, Culture, Sports, Science, and Technology "to intensify discussions with other countries." The ILC could start operations in the mid 2030s, he says. That would have the advantage of overlapping with the high-luminosity LHC, which is expected to start up in 2027 and run through the end of the next decade.

Another Higgs factory option is the Compact Linear Collider (CLIC), a long-standing project at CERN (see PHYSICS TODAY, May 2012, page 27). With CLIC, electrons and positrons could be accelerated to collide at 380 GeV over a comparatively short distance of 11 km, and the design could be extended to 3 TeV. Although the strategy doesn't favor it, CLIC could be revived if the 100 km FCC "ends up with a thumbs down," Bethke says.

The other Higgs factory candidates are circular. CERN is contemplating an electron–positron collider, the FCC-ee, in the same 100 km tunnel that would later house the FCC hadron–hadron collider. China has a similar proposal, the 100 km Circular Electron Positron Collider (CEPC; see "China plans a Higgs factory," PHYSICS TODAY online, 17 December 2018). It too could later be converted into a hadron collider. Synchrotron radiation limits a circular accelerator's energy reach; the foreseen energy range for both machines' electron–positron incarnation is roughly 90–350 GeV. At those energies, circular machines offer higher luminosity than do linear ones and could accommodate multiple detectors.

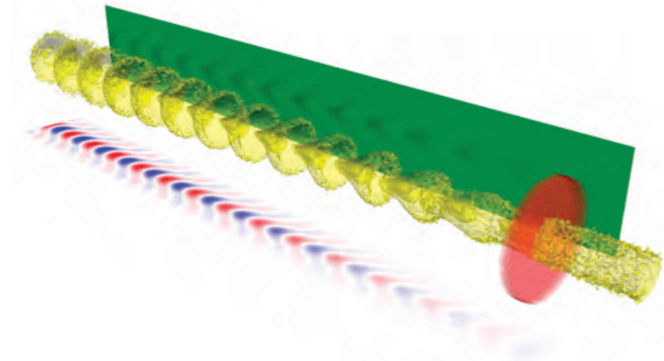
Even if the ILC goes ahead, it could still be worth building the FCC-ee. "They can be complementary," says Abramowicz. The ILC would start earlier, and if it gets extended to 500 GeV, Higgs self-coupling could be studied. With the FCC-ee, "we can go to lower energies and do extremely precise measurements of other standard model couplings," she says. "And we could obtain huge statistics to study the conservation or violation of flavor symmetry." But given the enormous costs involved, building two electron–positron colliders is questionable, says Klein, and many particle physicists say that if the ILC goes ahead, Europe should go directly to a pure hadron FCC.

The energy frontier

The ultimate goal is an FCC that produces hadron collisions, says Sijbrand de Jong of Radboud University in Nijmegen, the Netherlands. "By extending the energy frontier, we go into a region of energy that no human has explored. To convince others, we will have to come up with more tangible bread-and-butter goals for the funding agencies. There is a

S. BETHKE, MPP

DOZENS OF PARTICLE PHYSICISTS MET IN BAD HONNEF, GERMANY, in January to hammer out the details of the 2020 European Strategy for Particle Physics. Halina Abramowicz (center, front) chaired the strategy group. Just behind Abramowicz are CERN director general Fabiola Gianotti (plaid scarf) and CERN Council president Ursula Bassler (blue sweater). Several people quoted in this story are in the front row: Siegfried Bethke (second from left), Jorgen D'Hondt (fifth from left), and Leonid Rivkin (far right).



JOHN FARMER

THIS SIMULATION shows a laser pulse (red) flying to the right through a neutral gas and creating a plasma. Behind it, injected protons (yellow) accelerated by the Super Proton Synchrotron at CERN travel through the plasma and form bunches that drive the plasma wake that accelerates injected electrons (not shown). The European Strategy for Particle Physics recommends that physicists continue to advance plasma wakefield acceleration. In addition to their potential use in future colliders, plasma wakefield accelerators based on lasers could be incorporated into tabletop accelerators in hospitals.

lot of potential, but not everything is sorted out.”

Building the FCC at CERN will entail passing its 100 km tunnel under Lake Geneva (see “CERN considers a 100 TeV circular hadron collider,” PHYSICS TODAY online, 5 February 2019). Determining whether the geology is suitable and permits could be obtained is part of the feasibility study. The project would also pass under private lands and include aboveground shafts and a surface power grid, so local public acceptance is crucial. Environmental issues, such as where to put the extracted dirt, also come into play.

“The big elephant is power consumption,” says Jorgen D’Hondt of Vrije University Brussels and a strategy coordinator. A Higgs factory will consume up to three times as much energy as the current LHC, he says. “If you are part of the problem, you should be part of the solution. We need to develop better and alternative technologies to use less power.” In addition, he says, waste heat from the accelerators and cooling and computing systems can be more effectively reused.

The main technological hurdle is to produce superconducting 16 tesla magnets—nearly double the strength of the existing LHC magnets—to steer and focus the circulating protons. Work is underway using niobium–tin, but the material’s brittleness makes it tricky to form into the superconducting wires needed to make an electromagnet. So far, 11 T magnets are set to be tested in the high-luminosity LHC, and scientists have achieved stronger demonstrator magnets. “We estimate it will take 20–25 years to make 16–20 T niobium–tin magnets,” says Bethke. Scientists are also working on high-temperature superconducting magnets.

Finally, a robust financial plan with committed partners has to be in place before the European particle-physics community could recommend going ahead with the FCC. “The FCC will be far outside the CERN budget,” Bethke says. “We would need substantial additional contributions from the member states, and also from Japan, the US, Russia, and others.”

It could be difficult to justify building the FCC if China goes ahead with a similar machine. The CEPC “would be a big, ambitious project, even for China,” says the University of Chicago’s Kim, who is on the CEPC International Advisory Committee. “It could be an object of pride.” The collider is among several big international projects that may be considered in the country’s next five-year funding plan. To succeed, says Kim, “the project would need significant technical contributions and scientific involvement from the international community. It would also need financial contributions from other regions.”

The rising tensions between the US and China could make it difficult to implement a new project in which the two would be involved. But if China goes ahead with a Higgs factory, it could change Europe’s plans, says D’Hondt. “Whatever new collider emerges across the globe would trigger a reevaluation of our strategy.”

The European strategy recommends that the particle-physics community continue research on alternate acceleration techniques. “We have made up our mind for the FCC program,” says D’Hondt, “but to be prudent, we continue to investigate other options.” Chiefly, muon colliders, plasma wakefield accelerators, and CLIC; all three are widely viewed as approaches that could be adopted in the more distant future or serve as backups should the FCC not work out.

The US particle-physics community considered muon colliders, and then in 2014 largely abandoned them. But the idea is gaining renewed attention worldwide. The advantage of colliding muons, says D’Hondt, “is that you could access comparable physics with 10–15 TeV muon collisions as with 100 TeV proton collisions.” The difficulty with muons, he explains, is their short lifetime. “You have to put them in bunches immediately and collide them.” In addition, high-intensity, high-energy muon beams generate neutrinos and other particles at fluxes that, through interactions with atoms that create hadron showers, can be hazardous to experimental equipment and humans.

In plasma wakefield acceleration, an

incident laser beam or other source makes the electrons in an ionized gas oscillate along the beam's transverse field. (See PHYSICS TODAY, January 2015, page 11.) That creates a bubble of missing charge that moves near the speed of light; particles injected into the bubble experience a strong accelerating field. "It's useful for acceleration because of the strong electric field, but positrons still look difficult," says Allen Caldwell of the Max Planck Institute for Physics. And reaching high luminosities along with high energies is a challenge, he says. "It's not shovel ready."

Caldwell is pleased that the strategy mentions plasma wakefield and other alternative acceleration approaches and that it acknowledges the importance of neutrino physics, dark-matter searches, and other research directions. Still, he says, "the big colliders are front and center in the strategy. It's possible that noncollider particle physicists feel left out." Caldwell continues, "The strategy is conservative for my taste. This is a program that will go on for decades if it's realized. Is it enough to capture the imaginations of the next generations? Is it bold enough to keep the attention and focus of the particle-physics community for the next 50 or more years?"

Sustainability and society

A first in the 2020 strategy is a statement on environmental and societal impacts. "We are big energy consumers, and we want to think about how to recirculate energy," says Abramowicz. "We use gases that are not good for the planet, and we want to avoid liquids that if spilled pollute the environment." The strategy document encourages reducing travel. It also says that "a detailed plan for the minimisation of environmental impact and for the saving and re-use of energy should be part of the approval process for any major project."

Leonid Rivkin, of the Paul Scherrer Institute in Switzerland and one of the strategy coordinators, notes that the goals stretch out until at least 2080. That's why the strategy puts a strong emphasis on accelerator R&D and the associated educational push, he says. The strategy calls for particle physicists to "work with educators and relevant authorities to explore the adoption of basic knowledge of elementary particles and their interactions in the regular school curriculum." Says de Jong, "People should know about the structure of matter at a level deeper than the proton. It's part of mainstream culture, and it touches on societal issues—nuclear power, fusion power, medical diagnostic and treatment tools. These are relevant to people's daily lives."

More immediate is the question of whether and how COVID-19 may affect implementing the strategy. CERN's annual budget of about 1.3 billion Swiss francs ($1.4 billion) comes from member nations and is fixed through a legal convention. So far it appears stable. The pandemic and associated economic woes are a great concern, says Rivkin. "We hope that the member states realize the importance of particle physics and science to humanity to develop technologies and address societal problems. We motivate our bright young people and challenge them to come up with ideas for the future." It's important to fund the smaller national labs, too, he adds. "If you let the roots die, the tree will also die."

Toni Feder PT

Dawn of the TOPOLOGICAL AGE?

Arthur P. Ramirez
and Brian Skinner

Nontrivial electron band structures may enable a new generation of functional materials.

ISTOCK.COM/ALEXLMX

Arthur Ramirez is a professor of physics at the University of California, Santa Cruz. **Brian Skinner** is an assistant professor of physics at the Ohio State University.

Historians often label epochs of human history according to their material technologies—the bronze age, the iron age, and, most recently, the silicon age. From a physicist's perspective, the silicon age began with the theory, experiment, and device prototyping of a new type of material: the semiconductor. Although semiconductors had been known since the late 1800s as materials with unusual sensitivities to light, direction of current flow, and method of synthesis, not until the early 1930s did Alan Wilson make the radical proposal to describe their conduction in terms of the filling of their electronic bands.[1]

At the time, the concept of energy bands was firmly established, but electron conduction mechanisms were not clear. In the view of Felix Bloch, whose theoretical work on atomic crystals underlies the modern understanding of conduction, metals and insulators were just opposite limits of a continuous electron itinerancy. Wilson instead proposed that band filling is the control parameter: A filled valence band allows conduction only through electrons that are excited across an energy gap to another band, whereas electrons in partially filled bands can readily conduct by scattering into nearby states.

Wilson and others recognized that bandgaps were often controlled by impurities, but how impurities functioned was poorly understood. (Wilson incorrectly speculated that silicon in its purest state was a metal.) The 15 years following Wilson's proposal witnessed breakthroughs in purifying and controlling dopants in the elemental semiconductors silicon and germanium. Those advances eventually enabled the discovery of transistor action at Bell Labs in 1947. A surprise came, however, during the transistor patent preparation: The basic idea underlying the field-effect transistor had already been patented in 1930 by Julius Lilienfeld, an Austro-Hungarian physicist who had emigrated to the US in 1921.

For semiconductors, the path from theoretical understanding to device implementation was neither linear nor easily predicted. Topological materials seem to be following a similar trajectory. We have theoretical understanding and many ideas for novel devices, but ongoing materials development suggests the tantalizing possibility of our being at the dawn of a topological age. Here, we describe what it means for materials to be topological and how topology raises the prospect of revolutionary new devices.

Symmetry and invariance

Characterizing phases of matter by their symmetries is a central paradigm of physics. A magnet differs from inert iron because its internal magnetic moments consistently point in a particular direction rather than being isotropic. Similarly, a solid is different from a fluid because its atoms reside in fixed locations rather than moving freely. That prescription for understanding states of matter is usually referred to as the Landau paradigm.[2]

Over the past decade, however, awareness has grown that there is more to matter than the Landau paradigm.

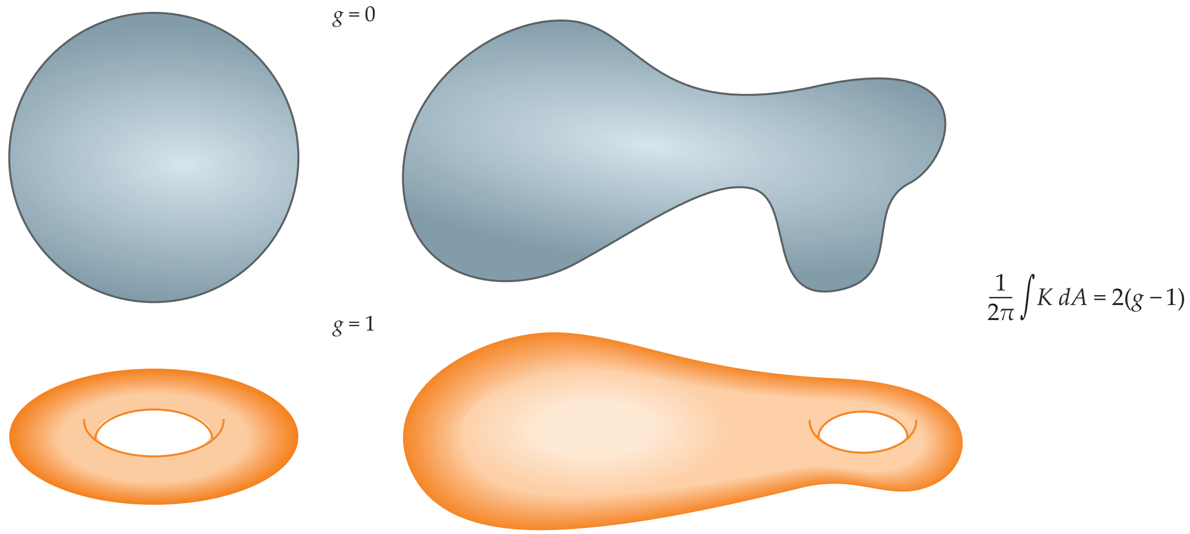


FIGURE 1. A TOPOLOGICAL INVARIANT is a property of a geometric shape that does not change when the shape is stretched or distorted. One such invariant is the genus g, which is given by the number of holes in the surface and is related to the integral of the Gaussian curvature K over the surface of the shape. Shapes with no holes in them ($g = 0$) all give the same value of this integral, as do shapes with one hole in them ($g = 1$). (Image by Donna Padian.)

Researchers are uncovering an ever-larger class of materials for which answers to basic questions, such as whether the material conducts electricity, depend not on local symmetries but on nonlocal properties called topological invariants. In much the same way that one cannot tell whether a coiled rope will form a knot when pulled tight unless one examines the full length of the rope, the electronic properties of a topological material can be determined only by examining the complete set of states in an electronic band. That nonlocality confers tremendous potential on topological materials: If a property is not defined locally, then it cannot be destabilized by local defects or fluctuations. The topological age thus promises a class of materials with unusually robust properties.

The notion of a topological invariant comes from the mathematical subfield of topology, which concerns those properties of geometric objects that are conserved under continuous deformations. The most famous such property is the genus g, an integer that counts the number of holes in a three-dimensional shape. (So $g = 0$ for a sphere, $g = 1$ for a donut, and $g = 3$ for a pretzel.) The genus is defined through the Gauss–Bonnet theorem, which states that the integral of Gaussian curvature K over the surface S of an object is quantized:

$$\frac{1}{2\pi}\int_S K\,dA = n.$$

Here n is an integer related to the genus by $n = 2(1-g)$. For example, consider a sphere with radius R. The curvature $K = 1/R^2$ is a constant, so the integral over the surface area $A = 4\pi R^2$ gives $n = 2$. That's consistent with $g = 0$, or an object with no holes in it.

The remarkable implication of the Gauss–Bonnet theorem is that if one stretches the sphere so that some parts of the surface become more curved and other parts become flatter, the integer n remains unchanged—it is topologically invariant (see figure 1). Much of the recent excitement surrounding topological electronics originates from the prospect of finding similarly invariant physical properties of electronic systems. Such a property would necessarily be robust to small perturbations or defects because integers cannot change continuously.

Topological electrons

In an isolated atom, electrons occupy discrete quantum energy levels, or orbitals. But when many atoms are arranged in a crystal lattice, the electron wavefunctions from neighboring atoms hybridize with each other and the orbitals broaden into bands of states, each having a range of energies. A state in a band describes an electron that is shared among many atoms, and the electron's wavefunction depends on the momentum $\mathbf{p}$ with which the electron hops from one atom to another. The wavefunction can be written as the product of two pieces: a plane wave that describes a free electron and a Bloch function $u_{\mathbf{p}}(\mathbf{r})$ that repeats periodically for each of the crystal's identical unit cells.[3] The Bloch function describes how the electron is affected by atomic nuclei in the unit cell, and as we will see, it contains information about the topology of the electron band.

The electron momentum $\mathbf{p}$ can have only certain restricted values. In particular, since $\mathbf{p}$ describes hopping between neighboring crystal lattice sites, the de Broglie wavelength $\lambda = 2\pi\hbar/p$ associated with the wavefunction cannot be shorter than the distance between neighboring unit cells of the crystal. The momentum in any given direction thus has a maximum possible magnitude. The space of allowable momenta is called the Brillouin zone, and its shape depends on the arrangement of atoms in the crystal.

In a discussion of an electron band's topology, the Brillouin zone plays the role of a geometric space. Closed surfaces in the Brillouin zone can be likened to geometric shapes that have an integer-valued index akin to the genus. Importantly, the Brillouin zone has effective periodic boundary conditions: Exactly opposite points on the zone's boundary are equivalent since they correspond to the same state with minimal de Broglie wavelength.

Constructing an analogue of the Gauss–Bonnet theorem for electron bands requires an analogue of curvature. As it turns out, that analogue arises from the properties of $u_{\mathbf{p}}(\mathbf{r})$. To see it, first consider the question: For a given momentum $\mathbf{p}$, where is the electron wavefunction centered in the unit cell? That question can be answered by calculating the expectation value of the position operator in the unit cell. The result is the quantity **X,** which is called the Berry connection. It can be thought of as the momentum-dependent center of the electron wavefunction in real space (see figure 2).

One caveat to the analogy with the Gauss–Bonnet theorem is that **X** is not precisely defined because its definition is not gauge invariant. The Bloch functions are defined only up to an overall phase that can depend on momentum. Thus the Berry connection is like the vector potential in problems with magnetic fields in that only its curl has a physical meaning. (We will show below that the analogy with magnetic fields runs deeper.)

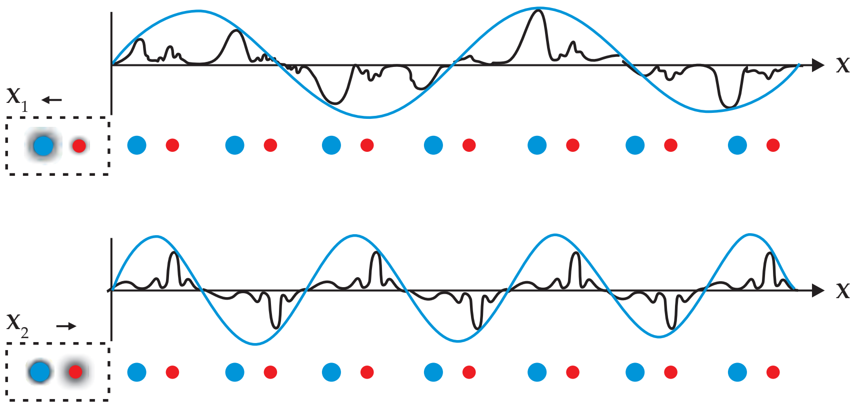


FIGURE 2. IN A CRYSTAL, AN ELECTRON STATE (black curve) is described by a slowly oscillating plane wave (blue curve) whose wavelength corresponds to the electron momentum that is modulated by a periodic Bloch function describing the electron's attraction to the atoms (red and blue circles) in the crystal's repeating unit cell. The electron probability density is shared among the atoms in the unit cell, as indicated by the shaded black areas in the outlined unit cells. The Berry connection **X**, shown above those unit cells, is a vector that can be thought of as the center of the state's density distribution. (Image by Donna Padian.)

Imagine now the hypothetical process of accelerating and then decelerating an electron such that the electron traces a path P in momentum space before returning to its initial momentum. The electron's initial and final states are identical except for a possible overall phase factor. That phase is an example of a Berry phase (see the article by Michael Berry, PHYSICS TODAY, December 1990, page 34) and it's given by

$$\gamma_P = \frac{1}{\hbar}\int_P \mathbf{X}(\mathbf{p}) \cdot d\mathbf{p}.$$

The accumulation of a Berry phase is analogous to the phase shift a particle traversing a path in position space experiences, which is equal to the number of wavelengths in the path multiplied by 2π. Taking the same path in momentum space in the clockwise and counterclockwise directions leads to opposite signs for γ_P because it reverses the direction of momentum change $d\mathbf{p}$.

The Berry phase becomes particularly instructive if we consider how it behaves for paths in a closed two-dimensional momentum space, such as the Brillouin zone of a 2D system, which is effectively closed because opposite edges of the zone are equivalent. Consider the green path shown in figure 3, which traces the boundary of a 2D Brillouin zone. Traversing the path in the clockwise direction yields a Berry phase $\gamma_{\rm BZ}$, whereas the counterclockwise direction gives $-\gamma_{\rm BZ}$. But opposite edges of the zone boundary describe physically equivalent states, so the clockwise and counterclockwise paths must produce equivalent changes to the wavefunction. That only happens if $\gamma_{\rm BZ}$ is either 0 or an integer multiple of 2π, leaving the wavefunction unchanged.

The condition on the Berry phase can be reformulated using Stokes's theorem to change the contour integral along the Brillouin zone boundary into a surface integral across the entire Brillouin zone. That procedure gives

$$\frac{1}{2\pi\hbar^2}\int_{\rm BZ} \Omega\, d^2p = C,$$

where Ω is the out-of-plane component of the Berry curvature, $\nabla \times \mathbf{X}$, and C is an integer known as the Chern number.

Having a nonzero Chern number requires symmetry breaking. In particular, in the Brillouin zone of a system with nonzero Chern number, the momenta $\mathbf{p}$ and $-\mathbf{p}$ are not equivalent; they have different values of **X**. That difference requires the system to break the symmetry with respect to either inversion or time reversal. The former leaves electron states unchanged when their spatial coordinates are inverted, and the latter when electron trajectories are played backwards in time. Thus, the search for topological materials has largely focused on materials that break one of those two symmetries. As we show below, only systems with broken time-reversal symmetry can have a nonzero Chern number; breaking inversion symmetry alone is insufficient. However, the coupling between electron spin and momentum may allow up and down spin species to each have a nonzero Chern number, so long as the two spin-resolved Chern numbers sum to zero.

Implications of topology

As illustrated in figure 3, a nonzero Chern number implies a winding or self-rotation in the structure of the electron wavefunction. That self-rotation is associated with the electron's physical angular momentum. For example, imagine making a wavepacket using states from a particular region of momentum space. The electron's position in the unit cell is related to the momenta of the states in the wavepacket. That relationship implies that the wavepacket's angular momentum depends on the local Berry curvature, making the Berry curvature again like a magnetic field: It's created by a broken symmetry in the material itself, and it gives electrons an orbital angular momentum.

The analogy of Berry curvature to magnetic fields becomes clearer when one considers the effects of an applied electric field **E**, which accelerates the electron. If the electron's center location **X** has a nonzero curl as a function of momentum, then as the electron accelerates, **X** shifts in the transverse direction. That shifting is known as an anomalous velocity, and it resembles the drift experienced by an electron in crossed electric and magnetic fields: Applying an electric field in a particular direction causes an electron to drift in a direction transverse to both **E** and the momentum-dependent **Ω**, which acts like a magnetic field.

One of the most striking implications of the magnetic field analogy arises from the motion of electrons at a sample's boundary. If a conductor with no intrinsic Berry curvature is subjected to a magnetic field, electrons near the boundary perform skipping orbits—they essentially roll along the boundary in a direction defined by the magnetic field. The skipping orbits persist no matter how the boundary is shaped and provide a single conducting channel for current to flow through. In a 2D electron system with a magnetic field and sufficiently high electron mobility, the skipping orbits give rise to the celebrated quantum Hall effect, with a quantized electrical conductance whose value is universal. Similarly, the self-rotation implied by a nonzero Chern number guarantees the existence of traveling edge states. Two-dimensional materials with nonzero Chern numbers have the same universal conductance, even if no magnetic field is present.

The existence of a topological invariant for electron systems subjected to a magnetic field was first identified by David Thouless, Mahito Kohmoto, Peter Nightingale, and Marcel den Nijs and bears their initials (TKNN).[4] Their topological invariant accounts for a remarkable universality of the quantum Hall effect among different samples and materials (see the article by Joseph Avron, Daniel Osadchy, and Ruedi Seiler, PHYSICS TODAY, August 2003, page 38). In fact, the TKNN invariant has allowed the universal constant $e^2/\hbar$ to be measured to more than 12 digits and now forms the basis for the metrological standard of the kilogram.[5] (See the article by Wolfgang Ketterle and Alan Jamison, PHYSICS TODAY, May 2020, page 32.) The Chern number can be thought of as a generalization of the TKNN result: Every material has a particular integer Chern number defined in the absence of any applied field. Most familiar materials have $C = 0$; recognizing the possibility of 2D materials with nonzero Chern number was a seminal insight of the topological age.

In quantum Hall systems, a magnetic field breaks time-reversal symmetry by forcing electrons to turn in spiral trajectories with a particular handedness set by the magnetic field direction. Playing those spiral trajectories backwards in time without reversing the sign of the external magnetic field produces motion inconsistent with the Lorentz force law. But edge states can exist in a topological material, even one that preserves time-reversal symmetry, if it combines broken inversion symmetry with a strong coupling between electron momentum and spin. (See the article by Xiao-Liang Qi and Shou-Cheng Zhang, PHYSICS TODAY, January 2010, page 33.) In the simplest case, those ingredients allow the two electron spin states to have nonzero but opposite Chern numbers.

To see how that situation can arise, consider that under time reversal, a left-moving spin-up electron becomes right-moving and spin-down. Thus, a topological electron band can retain time-reversal symmetry if the bands for spin-up and spin-down electrons have opposite Chern number (figure 4). The locking of edge-state directions—the two spin species moving opposite each other—is called the quantum spin Hall effect; it was discovered experimentally in 2007 following its prediction in 2003 (see PHYSICS TODAY, January 2008, page 19).

Ultimately, the locking of spin and momentum in edge states

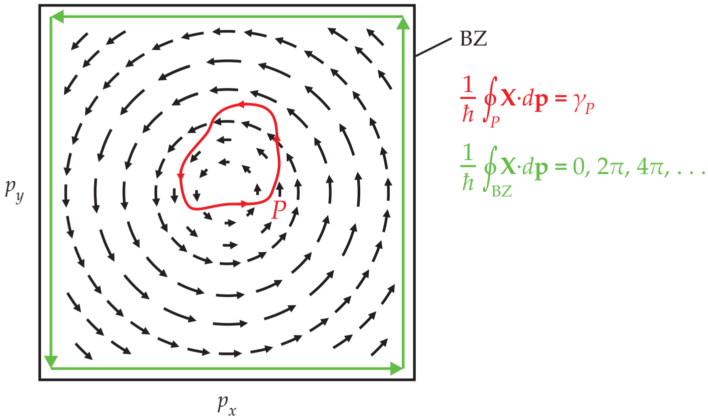


FIGURE 3. THE BRILLOUIN ZONE (BZ) is defined by the set of all possible momenta **p** for electrons in a crystal. This illustration shows the BZ for a two-dimensional crystal. The Berry connection **X** is a vector field (black arrows) in the BZ that indicates the electron wavefunction's center in the crystal's unit cell as a function of **p**. If an electron is accelerated and decelerated along some closed path *P* (red loop), its wavefunction acquires an overall phase γ_P whose sign depends on the direction of the path, clockwise or counterclockwise. But if that closed path runs along the BZ boundary (green loop), the phase must be a multiple of 2π. (Image by Donna Padian.)

arises from the microscopic spin–orbit coupling present in atomic orbitals. Spin–orbit coupling arises when a fast-moving electron experiences a magnetic field in its reference frame from the electrostatic potential of a nucleus. In the quantum spin Hall effect, strong spin–orbit coupling combines with broken inversion symmetry to produce a Berry curvature and a finite Chern number for each spin, even though no magnetic field is applied.

Topological bands in 3D

So far we have discussed only one example of a topological invariant: the Chern number in a 2D band that gives rise to edge states much like those in the quantum Hall effect. But 3D materials can also have electrical properties that are protected by a topological invariant. Those materials include topological insulators, which are usually narrow-bandgap semiconductors with strong spin–orbit coupling.[6] In topological insulators, an electrically insulating interior coexists with 2D metallic surface states in which electron spins are locked perpendicular to their momenta.

The 2D Chern number can also be applied to understand 3D Weyl semimetals. Such materials have special points in momentum space where their so-called topological charge is concentrated. To see why, imagine defining an arbitrary closed surface S of momentum states in the 3D Brillouin zone of some material (figure 5). Applying the same arguments about the Berry phase in 2D leads to the conclusion that the Chern number associated with the surface must be quantized. In particular,

$$\frac{1}{2\pi\hbar^2}\int_S \mathbf{\Omega}\cdot d\mathbf{A} = C_S,$$

where C_S is an integer that depends on the chosen surface S and describes a flux through the surface.

Given that C_S is an integer and cannot change continuously,

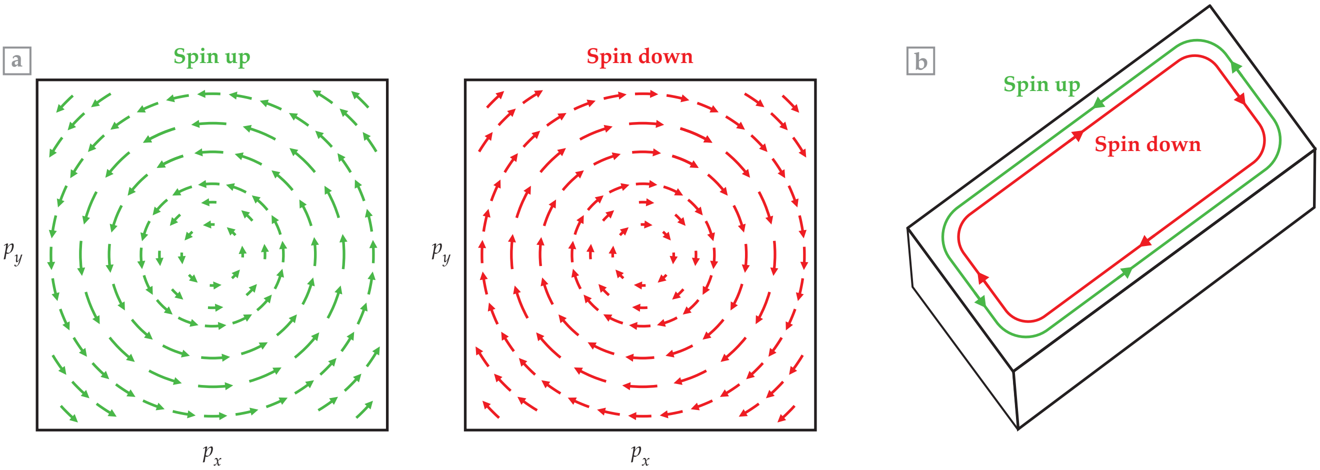


FIGURE 4. A QUANTUM SPIN HALL SYSTEM has equal and opposite Chern numbers, which describe electron wavefunction winding, for its two spin species. **(a)** The Berry connection **X**, shown here for a two-dimensional quantum spin Hall material, winds counterclockwise for up spins and clockwise for down spins, which gives Chern numbers of +1 and −1 respectively. The system remains symmetric under time reversal, which simultaneously changes **p** to −**p** and spin up to spin down. **(b)** The boundary of a quantum spin Hall material features edge states in which one spin species moves clockwise around the sample while the other moves counterclockwise. (Image by Donna Padian.)

slight distortions of the surface S cannot produce small changes to the integral. The only way for C_S to change is by a discontinuous jump, which happens when S is expanded to include a special point in momentum space known as a Weyl point. Weyl points are topological analogues to electric charges—they are monopoles of Berry flux—and the surface integral above mirrors Gauss's theorem. The points always come in pairs with opposite topological charges.

From a materials perspective, Weyl points arise when strong spin–orbit coupling causes two bands of electron states with different angular momentum to coincide in energy. The Weyl points correspond to the locations in momentum space where the two bands touch (see figure 5b) and the orbital character of the wavefunction changes abruptly. In the metals and semiconductors that make up most electronic technologies, such touching of bands is unusual. Typically, electronic bands cannot coincide in energy because of avoided crossing—the hybridization of degenerate quantum states into symmetric and antisymmetric combinations that have different energies.

In 1937, before the advent of modern topological band theory, Conyers Herring explained that two electron bands could meet—meaning they have the same energy—because of accidental degeneracies that prevent the two bands from hybridizing.[7] In that case, a perturbation that removes an accidental degeneracy can destroy the crossing and open a gap. In Weyl semimetals, however, the Weyl points are protected by the quantization of the Chern number. Only a sufficiently strong perturbation that brings two oppositely charged Weyl points together can destroy the degeneracy. Thus, Weyl semimetals are topologically protected gapless systems. Like other topological materials, they have intriguing surface states; in particular, their surfaces exhibit Fermi arcs, which are momentum states that connect the Weyl points.[8]

Identifying and classifying topological materials remains a challenge. Materials are often grouped based on their band structures; that grouping works well for semiconductors whose energy gaps largely determine functionality. But topological materials are defined not only by their energy spectra but also by the symmetries of their electron wavefunctions. Topological quantum chemistry aims to capture that complexity in order to characterize materials based on both symmetry criteria and conventional band structures.[9] The computationally intensive endeavor has unexpectedly found that an estimated 27% of all materials are topological in nature. Searches have uncovered not only new topological materials, but entirely new classes of them, such as nodal-line semimetals[10] in which two bands touch along a line rather than a point in momentum space, and higher-order 3D topological insulators[11] whose edge states exist only as lines or points along the hinges or corners of the crystal.

The interplay between topological electrons and acoustic or magnetic excitations is also a burgeoning field of study. Topological concepts can even apply to phonons and magnons themselves, which suggests a vast terrain of new materials is waiting to be explored.

Experimentally, the study of topological materials is progressing rapidly. New compounds and even whole classes of topological materials are routinely being discovered. In 2008 M. Zahid Hasan and coworkers at Princeton University first observed 3D topological insulating behavior[12] using angle-resolved photoemission spectroscopy (ARPES) in $Bi_{1-x}Sb_x$. Since then, many other topological insulators, Weyl semimetals, and nodal-line semimetals have been identified.

Technological prospects

Many materials derive their utility from their ability to either pass a current or prevent one from flowing. For example, the copper in a wire is useful because it allows electric current to flow freely, whereas the polymer encasing the wire stops the current from leaking out. Other materials pass or block heat currents, as do heat sinks on computer processors, or filter light, as do the frequency-selective lenses on protective sunglasses. From that perspective, the silicon age arose because semiconductors act as switchable valves for electrical current. We now know that pure silicon is a good insulator that only conducts if a gate voltage is applied to its surface.

New electronic materials usually have two performance

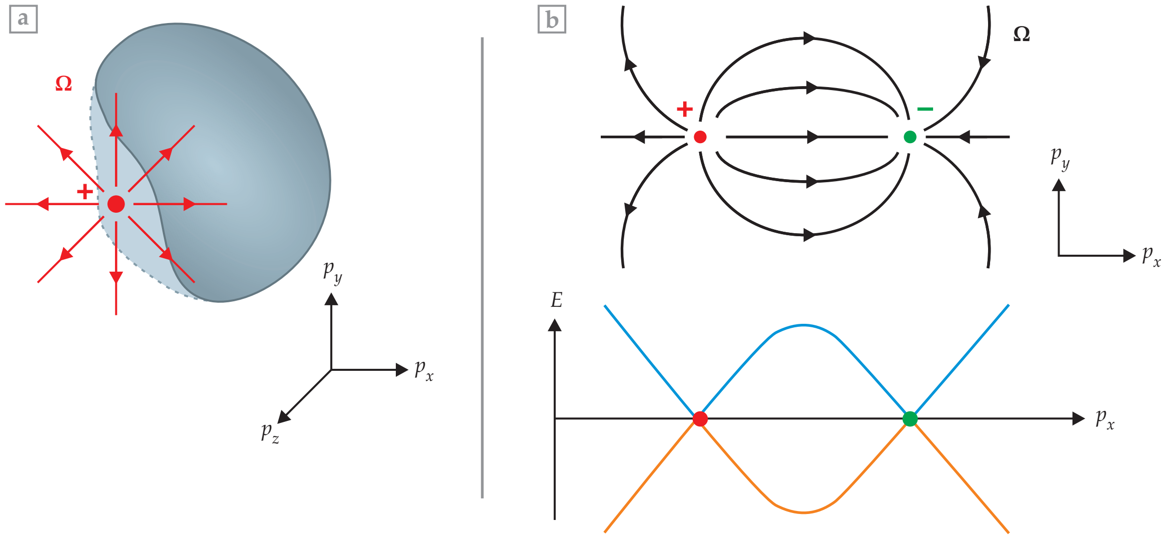


FIGURE 5. BERRY CURVATURE IN A WEYL SEMIMETAL stems from monopole sources, known as Weyl points, where two electron bands meet in momentum space. **(a)** The Berry curvature **Ω** can be depicted as a vector field (red arrows) emanating from or flowing into a Weyl point. A closed surface (gray) that does not enclose any Weyl points has Chern number $C = 0$. If that surface is expanded (dashed line) to contain a Weyl point, it abruptly attains $C = \pm 1$. **(b)** In a Weyl semimetal, Weyl points (red and green dots) come in pairs with opposite topological charge and are located at different momenta (top). At each Weyl point, two electron bands (blue and orange) meet in energy (bottom). (Image by Donna Padian.)

targets: filtering and sensitivity. The material should be able to selectively pass or block a generalized current in much the same way as silicon selectively transmits electrical current. It should also exhibit a strong response to some input, like the way silicon p–n junctions turn light into electricity. Topological materials offer the promise of truly new technologies in those areas. They are interesting filters because their Berry curvature is a kind of handedness that breaks the symmetry between clockwise and counterclockwise motion. Topological materials can therefore act like doorknobs that open when turned in the correct direction but block motion in the wrong direction.

One striking application is in spin filtering. As illustrated in figure 4, the edge states in topological insulators carry electrons with opposite spins in opposite directions. Such filtering is an essential ingredient for so-called spintronics, which aims to build electronic and computer technology based on currents of spin rather than charge.[13] The Berry curvature also implies that circularly polarized light would couple differently to the two electron species; that coupling could be used to create optical filters or logic circuits.[14]

Topological materials are unusually responsive to many kinds of applied fields because of their gapless, topologically protected electron bands. For example, the topological edge states associated with a nonzero Chern number could serve as dissipationless current-carrying channels. They have the potential to replace superconductors for some applications, and they may even function at room temperature.

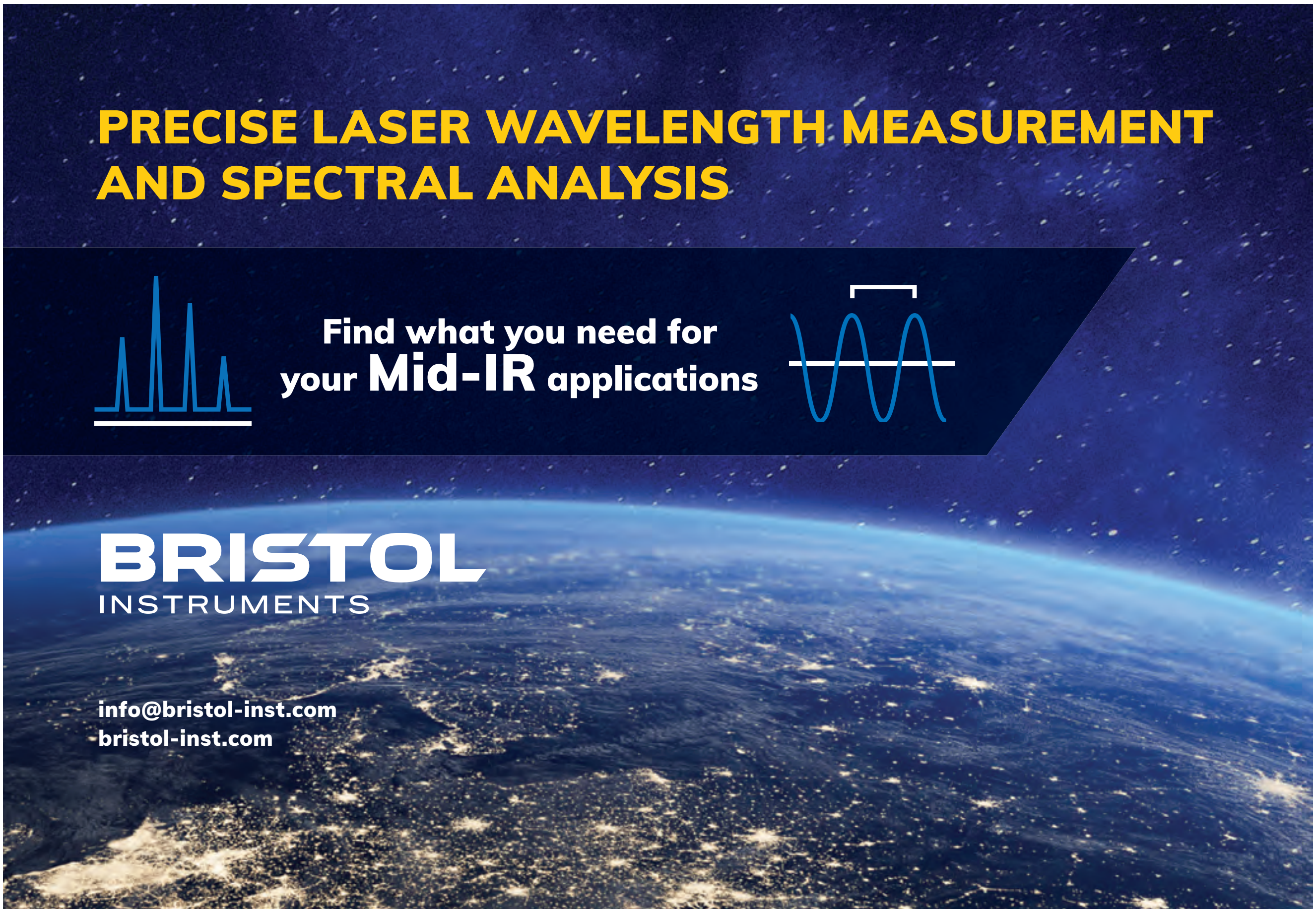
More generally, the protection of low-energy states in topological materials can be exploited in ways that would give them an advantage over conventional materials in which low-energy states are often distorted by disorder. The protection of a material's electron band structure can cause its electrons to exhibit enormous mobility, which results in each electron making an outsized contribution to a current being carried.[15] Weyl semimetals are extremely sensitive to light, which may lead to a new generation of photodetectors and night-vision goggles.[16] (See PHYSICS TODAY, July 2020, page 18.) Topological semimetals also exhibit an unprecedented thermoelectric effect, the ability to convert waste heat into useful electric power.[17] Additionally, topological electrons are unusually sensitive to magnetic fields. For example, in a magnetic field the quantum levels of the electron—its Landau levels—are widely spaced in energy, and applying a magnetic field along the current direction strongly reduces the material's electrical resistance, a phenomenon known as the chiral anomaly.[18]

Whether topological materials will revolutionize our current electronic technologies remains to be seen. But ideas from topology have clearly established themselves in materials physics, and they are here to stay. They have led to predictions and observations of new materials and phenomena. Who can tell whether they will come to define our current era?

REFERENCES

1. See M. Riordan, L. Hoddeson, *Crystal Fire: The Birth of the Information Age*, W. W. Norton (1997).
2. L. D. Landau, *Zh. Eksp. Teor. Fiz.* **7**, 19 (1937).
3. F. Bloch, *Z. Phys.* **52**, 555 (1929).
4. D. J. Thouless et al., *Phys. Rev. Lett.* **49**, 405 (1982).
5. K. von Klitzing, *Phys. Rev. Lett.* **122**, 200001 (2019).
6. C. Kane, J. Moore, *Phys. World* **24**(2), 32 (2011).
7. C. Herring, *Phys. Rev.* **52**, 365 (1937).
8. X. Wan et al., *Phys. Rev. B* **83**, 205101 (2011).
9. B. Bradlyn et al., *Nature* **547**, 298 (2017).
10. C. Fang et al., *Chin. Phys. B* **25**, 117106 (2016).
11. F. Schindler et al., *Sci. Adv.* **4**, eaat0346 (2018).
12. D. Hsieh et al., *Nature* **452**, 970 (2008).
13. I. Žutić, J. Fabian, S. Das Sarma, *Rev. Mod. Phys.* **76**, 323 (2004).
14. K. F. Mak, D. Xiao, J. Shan, *Nat. Photon.* **12**, 451 (2018).
15. T. Liang et al., *Nat. Mater.* **14**, 280 (2015).
16. C.-K. Chan et al., *Phys. Rev. B* **95**, 041104(R) (2017).
17. B. Skinner, L. Fu, *Sci. Adv.* **4**, eaat2621 (2018).
18. A. A. Burkov, *J. Phys.: Condens. Matter* **27**, 113201 (2015).

PT



TENURE-TRACK FACULTY POSITIONS IN PARTICLE PHYSICS AND COSMOLOGY

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

The Department of Physics invites applications for tenure-track faculty positions at the Assistant Professor level in experimental and theoretical physics, with specialty in the areas of **High Energy Theory and Cosmology, Particle Physics Experiment, and Observational Cosmology**.

Appointments at the rank of Associate Professor or above will be considered for candidates with an exceptional record of research excellence and academic leadership. The current faculty at The Hong Kong University of Science and Technology in particle physics and cosmology group include Professor Andy Cohen, Professor George Smoot, Professor Henry Tye, Professor Tao Liu, Professor Yi Wang and Professor Kirill Prokofiev. The department is growing its effort in particle physics and cosmology group by hiring five new faculty in theory and experiment. Further information about the Department can be found at http://physics.ust.hk.

Applicants must possess a PhD degree in physics or a related field. The successful candidates should have a strong track record of research. In addition to pursuing a vibrant research program, the candidates are expected to engage in effective teaching at the undergraduate and graduate levels.

Starting salary will be highly competitive and commensurate with qualifications and experience. Fringe benefits including medical and dental benefits, annual leave and housing benefits will be provided where applicable. Initial appointment will normally be on a three-year contract. A contract-end gratuity will be payable upon successful completion of contract.

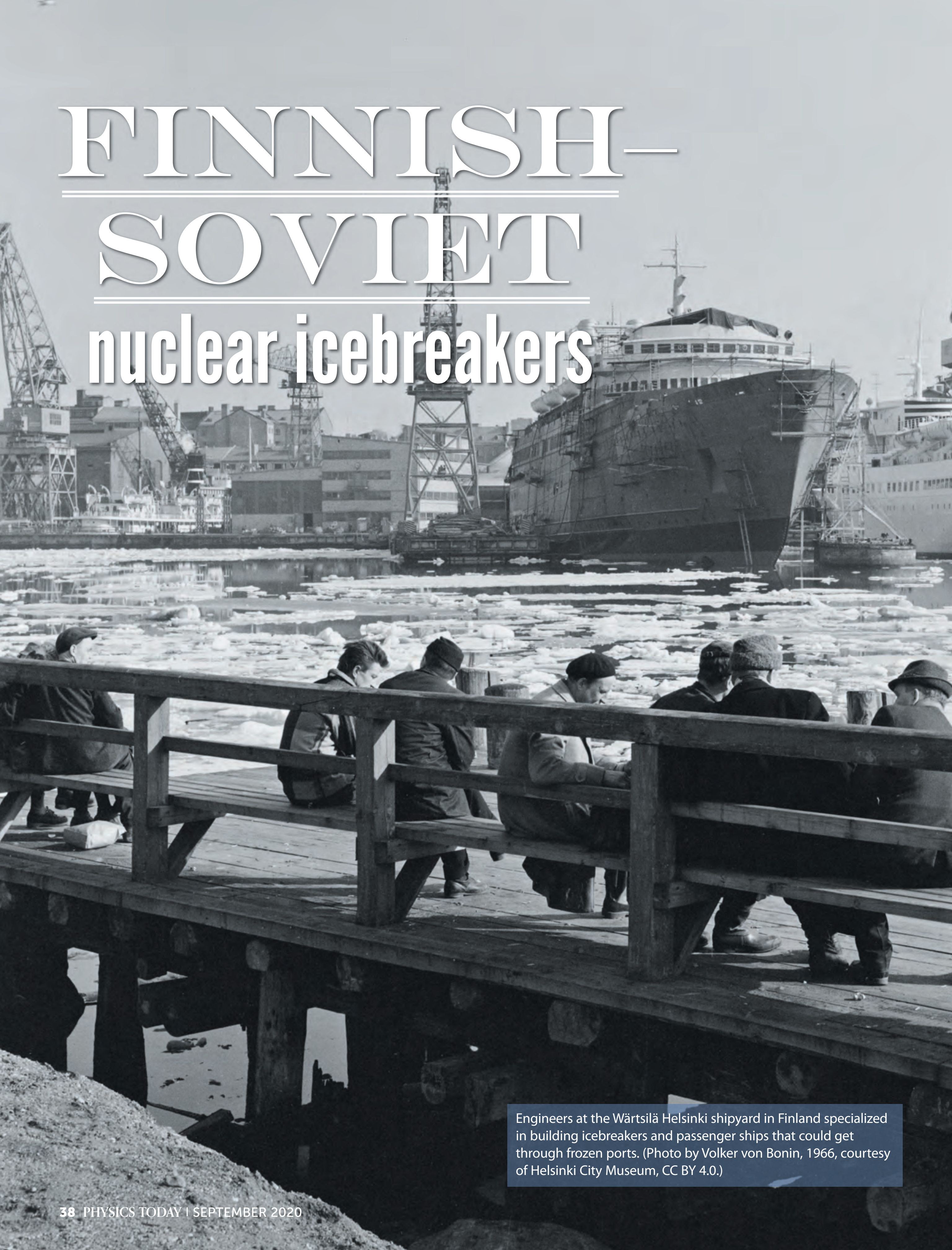
Application Procedure

Applicants should submit their applications along with CV, cover letter, complete publication list, research statement, teaching statement, and three reference letters. Separate applications should be submitted online for each position below:

High Energy Theory and Cosmology (PHYS1017H): https://academicjobsonline.org/ajo/jobs/16291
Particle Physics Experiment (PHYS1017P): https://academicjobsonline.org/ajo/jobs/16292
Observational Cosmology (PHYS1017C): https://academicjobsonline.org/ajo/jobs/16293

Screening of applications begins immediately, and will continue until the positions are filled.

FINNISH–SOVIET nuclear icebreakers



Engineers at the Wärtsilä Helsinki shipyard in Finland specialized in building icebreakers and passenger ships that could get through frozen ports. (Photo by Volker von Bonin, 1966, courtesy of Helsinki City Museum, CC BY 4.0.)

Saara Matala is a postdoctoral fellow at the Norwegian University of Science and Technology in Trondheim.

Saara Matala

During the Cold War, Eastern and Western manufacturers found good reasons to collaborate, even on a technology as sensitive as nuclear vessels.

In the spring of 1988, the shipbuilding company Wärtsilä in Helsinki, Finland, delivered the newest addition to the Soviet nuclear fleet: the icebreaker *Taymyr*. It was a rare occasion: During the Cold War, only two countries had built nuclear icebreakers, the gigantic vessels capable of opening up the Northern Sea Route (see figure 1). The first was the Soviet Union, which had expertise in nuclear propulsion and a long Arctic coastline. The second was Finland, a small country that had no expertise in building nuclear-powered vessels, no direct access to Arctic waters, and a lack of domestic demand for polar icebreakers.

Nuclear-powered ships carried heavy-weighted political meaning during the Cold War contest for ideological supremacy and national preeminence. They radiated state power, the capability of homegrown engineers, and national prowess.[1] That the Soviet Union, a nuclear superpower, ordered those imposing vessels from its small neighbor instead of building them itself was unprecedented.

The Finnish–Soviet nuclear icebreakers project was realized only because of the two countries' particular relationship and the politicized framework of East–West technology transfer. But the project diverges from the better-known Cold War stories of state-initiated strategic technology development. This story is about transnational cooperation rather than international competition, and its key actor was a private company rather than a government-backed enterprise. It's a story of how a small country harnessed technology to handle its strong neighbor rather than of how a big country used technology as a tool of power.[2]

Ships of national importance

In many ways, the Soviet Union and Finland were exact opposites. The Soviet Union was a vast and populous country, a nuclear superpower, and the leader of the socialist empire. Finland, on the other hand, was a sparsely populated capitalist country that had lost the Winter War of 1939–40 and the Continuation War of 1941–44 against its eastern neighbor. Unlike other small Eastern European countries, Finland was not occupied as a result of the war, and it maintained its democratic government. Yet its position in the no-man's-land of neutrality remained a sensitive issue throughout the Cold War.

The two countries did have at least one thing in common: interest in icebreakers. In Finland, foreign trade depended on the state icebreaker fleet during the winter months when all its ports froze. For the export-oriented country, uninterrupted maritime transportation was critical for national well-being, and surveying the coastline was essential for national sovereignty.[3] In the Soviet Union, icebreakers had been central to exploring and exploiting northern territories.[4]

The heavy-duty icebreakers were essential for supplying Siberian coastal traffic. After the onset of the Cold War, the strategic importance of the Arctic Ocean increased because it was the shortest route between the US and the Soviet Union. In both Finland and the Soviet Union, the public enthusiasm toward the giant state vessels went beyond their maritime function. They were widely recognized symbols of strength and power with well-known ship names. The icebreakers were technological monuments that contributed to Finland's national identity as a northern industrial country and to the narrative of Soviet socialist Arctic conquest.[5]

FIGURE 1. THE NORTHERN SEA ROUTE running along the Russian Arctic coast is an alternative route to travel between Europe and Asia but is usually only navigable with icebreaker vessels. (Image by iStock.com/Rainer Lesniewski.)

After World War II, the Soviet leaders felt it imperative to secure the strategic 1340-kilometer-long border with Finland. To that end, they needed it to have a trustworthy and compliant government, though not necessarily a communist one. The postwar priority of the Finnish government was to stay away from the superpower conflict and promote economic cooperation with East and West. Recognizing Soviet security interests, the Finnish leaders adopted a policy of neutrality that dominated foreign relations throughout the Cold War.

Finland had to balance its capitalist and democratic principles with pragmatism in establishing a stable relationship with the Soviet Union. The economic agreements that were launched after the war developed into a political relationship that demonstrated the diplomatic balance between domestic and international pressures.

For the Soviet Union, trading with neutral Finland benefited its economy and provided a potential extension for its domestic production. For Finland, economic cooperation with its socialist neighbor provided stability in domestic and international affairs and accounted for about a fifth of its foreign trade. The seemingly unsatisfiable Soviet demand for refined goods and machinery provided steady income for the growing Finnish urban working class. The arrangement demonstrated to Soviet leaders that a capitalist, friendly Finland was a more useful neighbor than an occupied, unstable satellite country. Even though most of Finland's foreign trade was directed to Western markets, the Finnish–Soviet agreements evolved into a symbol of peaceful coexistence between the communist and capitalist countries.[6]

After World War II, the Soviet Union desperately needed new tonnage. Having lost the war, Finland was forced to pay reparations worth $300 million, of which ships made up a considerable share. Most of them were relatively simple cargo ships and barges, but the large quantity required Finland to open new docks and to modernize its existing shipbuilding capacity. Once the reparation deliveries were completed in 1952, the Soviet market remained lucrative because many new yards weren't cost competitive with those in Western countries. From the Soviet point of view, the ship imports from Finland provided a way to rebuild its war-torn merchant fleet and bypass the export restrictions that limited purchases from more established Western shipbuilding countries.

The initiative for Finland to export icebreakers to the Soviet Union came from the private-owned company Wärtsilä in the early 1950s. It owned one of Finland's biggest shipyards in Helsinki, and it already had some experience building Finnish state icebreakers. Wärtsilä aspired to manufacture ice-going vessels that required specialized knowledge and would provide higher profit margins than simple cargo ships.

The rationale for the Eastern-oriented business strategy was clear. Demand for Baltic icebreakers was limited, but the Soviet Arctic offered almost endless market potential. The Soviet Union contracted with Wärtsilä in the early 1950s for the first of three *Kapitan*-class icebreakers, which were capable of cutting through ice in the Baltic Sea. From 1959 to 1969, Finland built a series of five *Moskva*-class polar icebreakers that were designed to cut through multiyear Arctic sea ice.

Early in the Cold War, icebreakers were strategic vessels that served military purposes. Originally, the Western bloc's Coordinating Committee for Multilateral Export Controls restricted the Soviets from buying such vessels from most Western countries. But the Finnish government understood the icebreakers to be civilian-service vessels, which did not conflict with the Finnish policy of military neutrality. For the Soviets, outsourcing part of their icebreaker building to Finland saved their advanced shipbuilding capacity for military vessels.

Promise of the atom

After World War II, nuclear technology encompassed a dual promise of life and death: fear of total destruction and hope for endless energy and well-being. When the dust had settled in Hiroshima and Nagasaki, scientists and politicians began looking to harness nuclear power for both military and civilian uses.

FIGURE 2. THE MODEL OF THE NUCLEAR ICEBREAKER *LENIN* was the center of attention at the Soviet industrial exhibition in Helsinki in 1959. (Photo by Väinö Kannisto, 1959, courtesy of Helsinki City Museum, CC BY 4.0.)

The widespread enthusiasm spanned a vast range of technical fantasies, from nuclear-powered trains and airplanes to pacemakers and typewriter-sized reactors for households. Global leaders recognized that the peaceful application of nuclear energy could also be used as propaganda to achieve cultural, political, economic, and technical goals.

From a technological perspective, naval architects who often struggled with the overall weight of a vessel found the lightness and durability of nuclear fuel especially attractive. Only a smidgen of uranium-235 was required to generate the same amount of energy as tons of diesel. A ship with a nuclear reactor on board seemed to promise a capacity to travel longer distances at lower cost and higher speed than vessels with alternative power sources.

The US military submarine *Nautilus* was the first vessel to realize those expectations. During sea trials in 1955, it traveled submerged at unprecedented speeds 10 times as far as any previous submarine. In the same year, President Dwight Eisenhower proposed building nuclear-powered merchant ships as part of his Atoms for Peace program. Four years later, Mamie Eisenhower christened the NS *Savannah*, a nuclear-powered cargo–passenger ship with a dual purpose of testing nuclear propulsion in a merchant ship and manifesting US expertise.

The Soviet Union launched the first nonmilitary nuclear surface vessel in 1957. The icebreaker *Lenin* had three pressurized-water reactors, which provided steam for turbines connected to electrical generators that drove the ship's propellers. *Lenin* was a considerable feat in the Cold War technology race and functioned as a prestigious symbol of socialist progress, as figure 2 illustrates.

In addition to serving civilian functions, both *Savannah* and *Lenin* demonstrated to the public the national technological competence and the potential uses of nuclear power. At a smaller scale but equally ambitious, other seafaring countries, including the nearby Scandinavian countries Sweden, Norway, and Denmark, examined the feasibility of nuclear merchant vessels, although they never launched any.

Whereas the Scandinavian countries entered the atomic age supported by their Western allies, Finland was still recovering from World War II and trying to adjust to the new international order. Early in the Cold War, Finland could not pursue nuclear power without endangering its fragile neutrality.

Despite not being a participant in the nuclear-technology race, Finland aspired to be a technologically advanced modern country. Its engineers followed foreign-technology journals, which provided them with copious reports of novel nuclear applications and new research projects. They read the same literature as their foreign colleagues and became enthusiastic about the possibilities of nuclear power.

Match made in heaven

The founding of the International Atomic Energy Agency in 1957 provided a way for Finland to obtain research reactors and opened a long-awaited opportunity for the Finnish engineers to satisfy their thirst for knowledge. When the Helsinki University of Technology started to offer nuclear-physics courses in 1959, among the students was Christian Landtman, a young technology manager from the Helsinki shipyard. Mathematics and theoretical physics astonished the practical mechanical engineer, and the laboratory exercises inspired his confidence in the applicability of nuclear physics.

The news of the Soviet nuclear icebreaker *Lenin* reached Finland in 1958 and roughly coincided with both the introduction of nuclear studies and Wärtsilä's entrance into the Soviet Arctic technology business. Landtman and the other young shipbuilding engineers at the Helsinki shipyard were struck by the concept of nuclear-powered polar icebreakers. Compared with the 22 000-shaft-horsepower (shp) diesel engines on *Moskva*, the reactors on *Lenin* could produce almost twice as much power for a year of nonstop travel. *Moskva* used 110 tons of diesel daily and required monthly refueling, which made it impractical in the most remote Arctic conditions, with long distances between the few ports. The long operational range of nuclear reactors and the lightness of their fuel met the challenges of the Arctic conditions, and the ice-strengthened hulls of polar icebreakers could protect the reactors.

After finishing an introductory course on nuclear physics, Landtman visited the construction sites of US icebreakers and the nuclear cargo ship *Savannah*, which was displayed proudly for foreign shipbuilders. US competence did not overawe the Finnish engineer. At a time of technological optimism, a nuclear icebreaker appeared feasible to the ambitious Landtman and to the shipyard of engineers who shared his confidence.

A paper icebreaker

Under the leadership of Landtman, Wärtsilä expanded its design functions and created a relatively large and skilled drawing bureau with expertise in tailor-made special vessels. At the turn of the 1960s, the bureau sketched a draft of a 50 000-shp nuclear icebreaker. In his presentation to German shipbuilders in 1961, Landtman recounted how Wärtsilä had installed in the hull of the icebreaker *Moskva* an intermediate form of the reactors used in *Savannah* and *Lenin*. He boasted that neither the space nor weight constraints had caused any problems.[7]

However, anyone familiar with Finnish winter navigation would recognize that a strong and expensive vessel would be overengineered for Baltic ice conditions. The only possible customer for a nuclear icebreaker was the Soviet Union. Eisenhower had recently vetoed domestic nuclear icebreaker plans because of the high costs and low priority. Other Arctic-bordering countries had such infrequent traffic that a heavy-duty nuclear-powered icebreaker was unnecessary.

In Moscow in 1961, Wärtsilä's managers presented the nuclear icebreaker project to their contact person at Morflot, the Soviet Ministry of Merchant Marine and principal operator of the Soviet Arctic fleet. Although Morflot was interested, the Finnish proposal was exceptional and undoubtedly expensive. It needed to be politically valuable to supersede other conventional ship orders of the Soviet planning economy. Being nuclear and an icebreaker was simply not enough.

Without any tested experience in nuclear propulsion, the Finnish shipyard could promise nothing new to the Soviets. The Baltic shipyard in Leningrad had already demonstrated the domestic capability in building *Lenin* for polar conditions. What the Finnish project could offer, though, was a channel to Western technology. The Soviet foreign-trade strategists expressed an interest in ordering a nuclear icebreaker, under the condition that it would be powered by a Western reactor. *Lenin*'s original reactor installations malfunctioned and were suspected of causing accidents.

Technology transfer from West to East had already been used by the Soviets to modernize their economy in the interwar period, and they emphasized its role in the economic strategies in the 1960s.[8] In the ideological contest for the future, the Soviet Union thought it pragmatic to take advantage of capitalist technological developments.

The Finnish shipyard executives sent inquiries for marine nuclear propulsion systems to all possible Western producers. Whereas the US and West German companies refused to sell nuclear technology that would end up in the Soviet Union, the UK, Sweden, and France appeared receptive to the idea. British and Swedish reactors eventually dropped out of the running, primarily because the buyer prioritized reactor reliability and did not want to pay for the R&D that remained to be done. That left France as the only possibility. Landtman was invited to join the visit of Finland's atomic energy advisory board to the French Atomic Energy Commission in 1965, which indicated the high political priority of the project.

In France, the Finnish request for a marine nuclear reactor was received positively. However, while the French were pondering their response, the Soviets abandoned their original idea and instead opted to use their own reactors for the nuclear fleet. Simultaneously, they announced their own nuclear icebreaker project. The *Arktika*-class concept had an astonishing 70 000 shp; the Finnish design utilized 50 000 shp. Despite the setback, Wärtsilä kept itself busy building seven conventionally powered polar icebreakers and five smaller river and harbor icebreakers for the Soviet Union.

But the Finnish nuclear icebreaker project survived. In 1970 a Finnish weekly magazine published a cover story on Landtman. He told his interviewer that the Helsinki shipyard's nuclear icebreaker project was on hold because it lacked a reactor, but he didn't mention that it also needed a hull, detailed plans, and a contract. Until the late 1970s, the Finnish nuclear icebreaker was little more than a fascinating idea and some technical drawings used to gauge Soviet interest. The project existed merely on paper, but it was still influential in making the Helsinki shipyard a serious actor in nuclear shipbuilding.

Repurposing the nuclear icebreaker

The 1970s marked a turning point in the public perception of nuclear technologies as postwar enthusiasm gradually gave way to realism. Project after project proved that nuclear propulsion was still more expensive and riskier in merchant shipping than oil-fired engines. The US's *Savannah* was retired in 1971 and never reached a break-even performance. The German nuclear cargo vessel *Otto Hahn* was completed in 1968 and was deactivated in 1979. The Japanese nuclear cargo ship *Mutsu* was blocked by protesters in her home port before the first test run. The early proponents of the nuclear merchant marine were hampered by real or imagined concerns about safety that prompted several ports to turn the ships away. In 1975 the *Guardian* opined, "The fact is that nuclear powered merchant ships cannot trade like other ships because they frighten the life out of governments and people."[9]

In Finland, the Helsinki shipyard resurrected its nuclear icebreaker project to help solve new problems. The structure of Finnish–Soviet trade had always been asymmetric: Finland imported crude oil and raw materials and exported machinery and refined products. From the 1970s on, the Soviets increasingly pushed to replace that colonial trading relationship. Although that aspiration barely affected the laws of supply and demand, joint technology projects between the two countries became politically invaluable. Collaborative projects demonstrated the Soviet Union's technical advancements, facilitated knowledge diffusion, and served as propaganda to trumpet their peaceful coexistence to the world. In its relationship with Finland, the Soviet Union wanted to be a technological pioneer, not a technology importer.

The Soviets attempted to increase the use of their own products in every ship they ordered from the Finnish shipyards. Soviet delegations were particularly eager to sell main engines because they made up a high share of a ship's total value. Most Finnish executives were equally eager to reject those sugges-

FIGURE 3: THE NUCLEAR ICEBREAKERS *Taymyr* (left) and *Vaygach* (right) were designed for travel in shallow coastal waters but are capable of breaking thick multiyear Arctic ice. (Photo courtesy of Aker Arctic.)

tions and provide business to their own engine producers. Instead, they tried to fill the quotas for Soviet content with simpler navigation systems, radio equipment, and some metal products, such as anchor chains. Although the Soviet Union had no legal way to force private enterprises to buy unwanted components, the Finnish shipyards realized they had to address the Soviet trade aspirations to remain on good terms.

When Wärtsilä proposed using the nuclear icebreaker as a potential Finnish–Soviet collaboration in the spring of 1977, shipbuilding engineers and politicians concerned with foreign trade approved. The Helsinki shipyard proposed designing and building the icebreaker, whereas the Soviet Union would construct and install the nuclear reactors. The combination of Finnish icebreaker expertise and Soviet nuclear technology was justified technically and economically. Wärtsilä had never planned to develop its own reactor technology. Outfitting an icebreaker with Soviet reactors was less a concession than a creative response to Soviet demands.

To differentiate the Finnish project from Soviet-built nuclear icebreakers, Wärtsilä specified it as a shallow-draft nuclear icebreaker. Because of their deep draft, *Arktika*-class icebreakers in the Soviet nuclear fleet were capable of reaching anywhere in the Arctic Ocean except the shallow coastal waters. To build an icebreaker that could go through two-meter-thick ice and navigate the shallow rivers and coast of Siberia, a company had to combine scientific understanding and technical experience in ice mechanics, naval architecture, and icebreaking. Wärtsilä had already demonstrated that specialized expertise in their other shallow-draft icebreakers.

At that stage, Wärtsilä's promotion of the nuclear icebreaker project resembled political lobbying more than technical marketing. It focused on persuading high-ranking Soviet politicians to order a Finnish vessel. In December 1978 Finnish president Urho Kekkonen returned from his visit to the Soviet Union with news that Morflot was willing to order two nuclear-powered icebreakers from Wärtsilä. Besides the nuclear reactors, the Soviet Union wanted to deliver special low-temperature hull steel originally developed for its submarines, turbines, and propellers. Those contributions raised the Soviet share of the project to about 13% of the cost and elevated the icebreaker project to showcase Finnish–Soviet scientific, technical, and industrial cooperation.

The two countries signed an official agreement in November 1980. For a project that had been launched almost 20 years earlier, the milestone was important, but the commercial negotiations were not completed until years later. The Soviets were known to announce valuable contracts during celebrations of important anniversaries, and that made the Finnish delegation excited on the eve of national holidays. Finally, in 1984, during the festive week commemorating the 1917 October Revolution, the Soviets agreed to sign the order for two nuclear icebreakers.

Four years later, in 1988, the first ship of the class was named *Taymyr*, after an old imperial Russian icebreaking steamer and the northernmost part of the Eurasian mainland. *Taymyr* is shown in figure 3 with its sister ship *Vaygach*, which was launched later the same year and completed after two years.

As the first nuclear-powered ship ever launched in Finland, *Taymyr* enjoyed undivided attention from the public. But the two nuclear ships were never radioactive in Finland. The icebreakers left Helsinki powered by temporary oil-fired boilers on the helicopter deck. Once the ship had docked in Leningrad, the Baltic shipyard installed the reactors. Having completed the project, Finland joined a group of just a few countries that built civilian nuclear vessels, even as most postwar nuclear dreams had withered away.

Ships carry political weight

Finnish shipbuilding engineers had an apolitical identity. Engaging in business with the Soviet Union was about generating profits. Yet as the history of the Finnish–Soviet nuclear icebreaker demonstrates, politics did matter for technological development. Even the Finnish engineers were aware of the political weight the nuclear ships carried. For them, Cold War politics was among the various means of persuasion the shipyard used to promote its product for its only potential customer.

At the beginning, the Soviet Union saw the Finnish nuclear icebreaker project as a channel for Western marine-reactor technology. At the end, the project motivated the industrial and technical cooperation between a private shipyard and Soviet governmental institutions. The Finnish state and the shipyard could display the nuclear icebreaker project as a symbol of advancement and a commitment to new technology and progress. The lure of the atom kept the project alive before the first contract was signed 20 years in, a long time for a profit-driven, privately owned company working on a nonmilitary project.

Technical qualities seemed to justify the Finnish–Soviet project: enough horsepower to break through thick polar ice, a long duration of fuel enabling independent operation in the Arctic, and a hull shape and lightness that together allowed the Finnish-built ships to navigate waters too shallow for other nuclear icebreakers. However, the first 20 years proved that closing the deal demanded more than technology. The project survived because Wärtsilä creatively and persistently aligned it with Finland's and the Soviet Union's political agendas. It was big enough to showcase technological cooperation across borders but did not challenge other Soviet nuclear projects. It was strategic enough to be prioritized above other projects and civilian enough not to endanger Finland's neutrality policy.

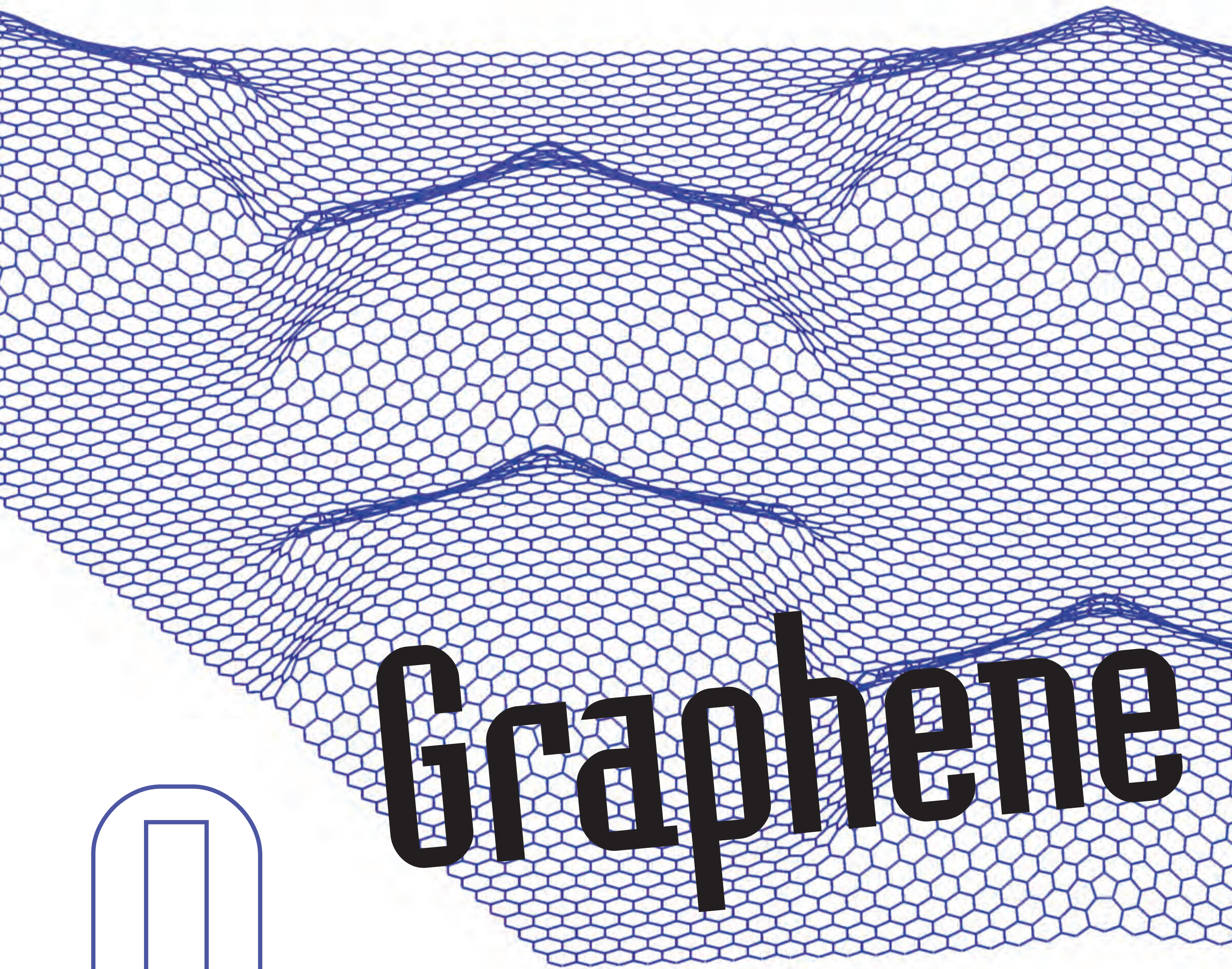
The Finnish–Soviet nuclear icebreaker project was political without being entirely governmental, and nuclear without being radioactive. But it was truly a child of the Cold War and would not have been possible without the politicized framework of the Finnish–Soviet technology trade. International crises such as a war always shape technological development by creating restrictions and opportunities. The full scale of the unexpected effects can often be evaluated only in hindsight.

This feature is adapted from "Flashy flagships of Cold War cooperation: The Finnish–Soviet nuclear icebreaker project," Technology and Culture ***60****, 347 (2019), with the permission of Johns Hopkins University Press.*

REFERENCES

1. G. Hecht, *Being Nuclear: Africans and the Global Uranium Trade*, MIT Press (2014).
2. S. Autio-Sarasmo, K. Miklóssy, eds., *Reassessing Cold War Europe*, Routledge (2011); N. Oreskes, in *Science and Technology in the Global Cold War*, N. Oreskes, J. Krige, eds., MIT Press (2014), p. 11.
3. S. Matala, A. Sahari, *Hist. Technol.* **33**, 220 (2017).
4. J. McCannon, *Red Arctic: Polar Exploration and the Myth of the North in the Soviet Union, 1932–1939*, Oxford U. Press (1998).
5. P. R. Josephson, *The Conquest of the Russian Arctic*, Harvard U. Press (2014).
6. S. Matala, "The Finlandisation of shipbuilding—Industrialisation, the state, and the disintegration of a Cold War shipbuilding system," PhD thesis, Aalto U. (2019).
7. C. Landtman, in *Jahrbuch der Schiffbautechnischen Gesellschaft*, vol. 55, Springer (1961), p. 142.
8. S. Autio-Sarasmo, in *Modernisation in Russia since 1900*, M. Kangaspuro, J. Smith, eds., Finnish Literature Society (2006), p. 104; K. E. Bailes, *Comp. Stud. Soc. Hist.* **23**, 421 (1981).
9. "Unwelcome nuclear ships," *Guardian*, 5 February 1975, p. 12.

PT



Graphene

A Japanese legend says that if you fold a thousand origami paper cranes then you will be granted one wish. If your wish is to tear those cranes and refold the paper into new shapes, you will be practicing a different papercraft: kirigami. In that craft, cutting and then folding a sheet lifts the two-dimensional material into a three-dimensional world. Kirigami is as old as paper itself, and it is gaining appreciation as a material-design tool with possible applications ranging from architecture to nanoscience.

You are probably familiar with the form of kirigami found in pop-up greeting cards, which have cuts in a sheet of paper that produce 3D shapes when the card is opened. But kirigami can also have a slightly different construction: Strips or regular sections of a 2D material are removed, and when the open edges are rejoined, the 2D sheet is forced into a 3D shape. You can try this with any sheet of paper you might have handy—although we discourage using this article.

Lattice kirigami takes that approach one step further by developing cut-and-fold rules appropriate for 2D lattices.[1] A lattice's discrete geometry determines the rules. First, material is removed by cutting lines or wedges. Then closing those opened regions leaves topological scars in the material, known as dislocations in the case of a line and disclinations in the case of a wedge. Finally, those defects deform 2D lattices into nonplanar 3D structures, the shapes of which generally depend on an interplay of the bending modulus, in-plane rigidity, and lattice anisotropy. Although that competition can be complicated, laterally stiff materials, which resist in-plane stretching, have a set of simple folding rules.[1]

Bastien Grosso is a PhD candidate in the department of materials at ETH Zürich. **Eugene Mele** is Christopher H. Browne Distinguished Professor of Physics at the University of Pennsylvania.

gets bent

Bastien F. Grosso and Eugene J. Mele

Two-dimensional nanomaterials are bending the rules of the papercraft known as kirigami.

Because lattice kirigami folding rules are purely geometrical, one would expect that they would be robust and apply across material platforms that span many length scales. But in practice, researchers have found that the rules change at the nanoscale, where smooth shapes replace sharp corners.

The composition rules for kirigami in continuous media—for example, sheets of paper—and discrete macroscopic 2D lattices, such as an array of foldable tiles, follow the simple universal governing principle: Bend but don't stretch. That idea fails for nanoscale materials, which surprisingly are found to relax by stretching even if the lattice is laterally very stiff. The result is three new bending rules for nanoscale kirigami. The rules predict, for example, the 3D shapes of graphene sheets that are inscribed with select patterns of topological lattice defects. For some shapes, graphene's electrons traveling along the curved surface undergo anomalous sideways deflections as though they were moving in a colossally strong magnetic field, much stronger than can be achieved with any terrestrial magnet.

Some assembly required

In both continuous and discrete media, the interplay between defects and the 2D sheets' bending and stretching energies is illustrated by the deflection of a flat surface into the third dimension by a disclination, such as the one shown in figure 1a. Any smoothly spatially varying out-of-plane displacement $f(r)$ costs a bending energy $(\nabla^2 f)^2$. If the sheet's shape has Gaussian curvature $K(r)$, defined as the determinant of the matrix $\partial_i \partial_j f(r)$, it also must have strain in the tangent surface—imagine trying to tightly wrap an orange in plastic wrap. The resulting strain field produces a phonon-mediated interaction between the Gaussian curvatures in widely separated regions[2] and incurs an additional stretching energy penalty.

If the medium is nearly inextensible—that is, if it bends easily but doesn't stretch—it tends toward shapes that minimize in-plane stress. But that's only possible if $K(r)$ can configure itself to locally compensate for the disclinations. For example, when a wedge is removed from a material, the leftover sheet forms a conical shape with the Gaussian curvature confined to the vertex, as shown in figure 1a. The shape eliminates the macroscopic stretching energy at the expense of a significant bending energy, which increases logarithmically with the size of the system.[3] Reconnecting the edges of the wedge spontaneously forces the deflection of the 2D surface into 3D, and in this instance doesn't require any stretching at all. Similar ideas appear in theories of stress focusing at vertices and ridges of crumpled elastically stiff sheets.[4]

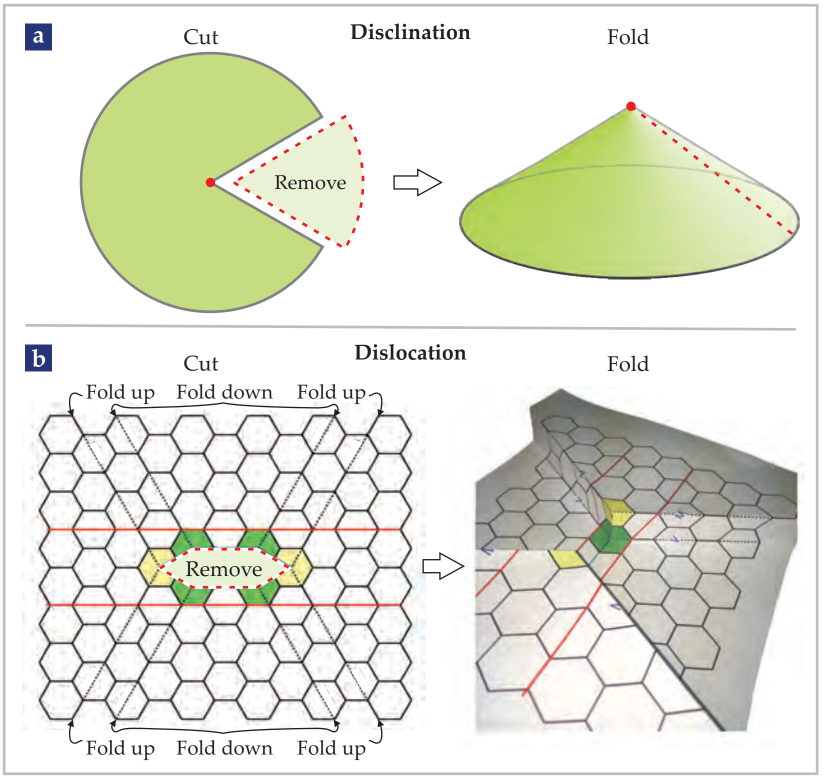


FIGURE 1. CUTS AND FOLDS PRODUCE KIRIGAMI STRUCTURES in a continuous material and a lattice. **(a)** When a wedge of material is removed from an unstretchable two-dimensional disk (left), the edges rejoin and force the disk into a conical shape (right), which localizes all its Gaussian curvature to the vertex and liberates bending energy into the curved surface. **(b)** With a discrete reformulation, kirigami in a lattice model takes a similar form. Material is removed from the green and yellow hexagons at the center, and the lattice folds along the dashed lines into the third dimension to seal the gap. The shape preserves the area of the incompressible lattice cells and localizes the Gaussian curvature to the corner vertices. (Adapted from ref. 1.)

Since a lattice dislocation can be viewed as a bound pair of disclinations, one might expect a related elastic compensation condition to hold. But the obvious mathematical approach of treating the dislocation-induced displacement as a superposition of Gaussian curvatures fails because the stability conditions are nonlinear. A continuous medium can't screen a dislocation with a Gaussian curvature field focused to a point dipole.[3] By abandoning the continuous picture in favor of a new set of rules adapted specifically for a discrete lattice, theorists can define a fully discretized variant that defines a family of shapes caused by dislocations. Randy Kamien at the University of Pennsylvania and his colleagues were the first to derive the variant for topological defects embedded in a 2D honeycomb lattice.[1] In that application, they treated the medium as a network of perfectly incompressible hexagonal units that are linked along easily bendable shared edges.

Kirigami on that honeycomb network takes a form like that shown in figure 1b. First, rows of hexagons are removed from the structure, then the material gap is eliminated by folding the sheet along the honeycomb edges (dashed lines). Closing the surface in that way has an appealing arts-and-crafts aesthetic, which can be appreciated by experimenting with a cut-and-fold paper model: One website (www.futurity.org/kirigami-hexagons-819442/) has a video showing step-by-step directions to build a paper model of the 3D finished product. The reconstituted lattice is a sharply folded structure with right angles and a ridge-and-plateau surface.

The surface is locally flat with the plateau's height set by the width of the cuts at the ends of the dislocation—that is, the lines from the center of a yellow hexagon to the center of a green hexagon in figure 1b. All of the material's residual Gaussian curvature lies at its sharp corner vertices, similar to the curvature confinement at the cone's vertex for the case of a disclination. The model's target shape nulls the stretching energy, and if the edges hinge perfectly and freely, it has no bending energy either. Composition rules for even more complex 3D surfaces show that all shapes can be reconstructed from just two possible kirigami incisions: cuts that pass through the hexagons' nodes or those that pass through the hexagons' centers. Similar in spirit to the continuum solution for an isolated disclination, the common directive for all the constructions is to permit extreme bending in order to avoid a more costly stretching energy.

Lattice kirigami replaces an intractable continuum-shape optimization problem with a discretized problem by prescribing folding rules that are local and essentially geometrical in character. Any mechanical structure with a set of similar composition rules could support similar sharply faceted kirigami surfaces. Those rules can be designed into macroscopic architectures by linking elementary rigid units with folding rules that guide the deflection of a flat surface into a 3D shape. Early researchers speculated that the same strategy could even control the shape and structure of the many robust microscopic

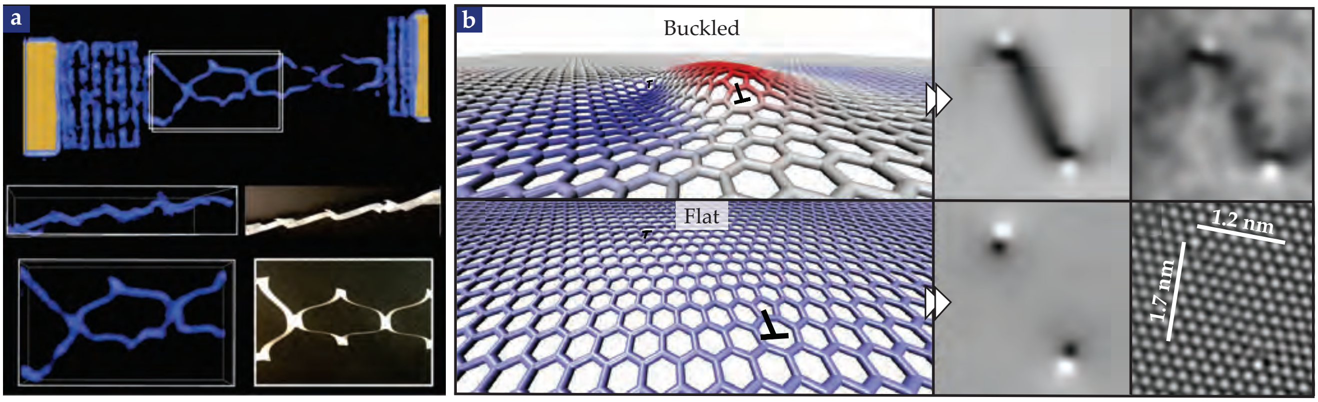


FIGURE 2. GRAPHENE DOES AND DOESN'T BEHAVE LIKE PAPER. (a) Lithographically patterned grooves in graphene produce a two-dimensional ultrastretchable spring. The spring's compliance depends on the bending modulus (bending) rather than the Young's modulus (stretching) because the spring buckles into the third dimension. The shape of the extended spring looks the same for paper (lower left) and graphene (lower right). (Adapted from ref. 5.) **(b)** Unlike in paper, a pair of dislocations (bottom row), or line cuts, in graphene are bridged by a lattice scar (top row) that develops a smooth humpback surface in simulated (center column) and experimental (right column) transmission electron scattering images. (Adapted from ref. 7.)

honeycomb lattices found in atomically thin 2D materials, such as graphene and transition metal dichalcogenides.

Rolling landscapes

Applications of lattice kirigami to nanomaterials challenge its two central tenets. First, in nanomaterials the bending modulus is nonzero, so the materials can't form sharply folded edges. Second, the system is compressible and thus can store energy in shear and compressive strains. In practice, the result is an inevitable tradeoff between bending and stretching deformations, and the shape is determined by the solution to a subtle optimization problem in a vastly larger parameter space.

For a single sheet of graphene, that tradeoff appears to resolve in favor of the same "bend but do not stretch" algorithm as a sheet of paper. In fact, graphene and paper are remarkably similar elastically. One method to quantify the stretchability of 2D material is the relative in-plane and out-of-plane stiffness, or the dimensionless Föppl–von Kármán number υK. For a square sheet of width L and thickness t, $\upsilon K = YL^2/\kappa \approx 10(L/t)^2$ in terms of the 2D Young's modulus Y and the effective bending rigidity κ.[5,6] Larger υK values indicate that a material is less willing to stretch. For a graphene sheet with width $L = 0.1$ µm, the estimate is that $\upsilon K \approx 10^7$, even larger than υK for an essentially inextensible sheet of paper.

That back-of-the-envelope estimate finds some experimental support. For example, a graphene sheet can be patterned into the same flat 2D spring as paper, see figure 2a. Paul McEuen and his group at Cornell University lithographically cut graphene with alternating grooves on a few-micron scale.[5] They found that springs with that design are ultrastretchable with a compliance determined by the bending rigidity instead of the Young's modulus. Bending rather than stretching occurs because the 2D spring buckles into the third dimension in a manner strikingly similar to the shape response from pulling on a macroscopic paper spring with the same design.

That graphene–paper correspondence is only partially reassuring. Crucially, the nanosprings are ultrastretchable only because they contain open perforations that allow the system to bend and not stretch. As a result, the springs lack the reconnected surfaces that contain topological defects, which are the raison d'être of lattice kirigami. A more relevant point of comparison is the fully bonded defect structure of the scars that form on dislocations produced in graphene grown by chemical vapor deposition.[7] There a very different picture emerges. Even in the extremely stiff limit with $\upsilon K > 10^5$, the scars form smoothly elevated humpbacked surfaces, as shown in figure 2b, that connect dislocations. Similar to the nanosprings, those scarred graphene structures are laterally very stiff. Their 3D surfaces, however, are not faceted but smooth, the Gaussian curvature is not focused but distributed, and the nearly unstretchable sheet is stretched.

Theoretical investigations lead to similar conclusions. Several years ago we (Grosso and Mele) carried out large-scale atomistic simulations to study the structures of dislocated graphene sheets that start in faceted lattice-kirigami structures and are then relaxed using accelerated molecular mechanics, a computer code that efficiently solves the equations of motion for many atoms simultaneously.[8,9] In all cases the structures revert to shapes that are warped and softly undulating, as shown in figure 3. Ridge-and-plateau kirigami is replaced by rolling landscapes evocative of the English countryside outside Loughborough University, where the work was carried out. Other researchers obtained similar results for the effect of disclinations in 2D disks of a phosphorus allotrope known as phosphorene.[10] The defected disks relax into domed 3D shapes that also redistribute their Gaussian curvature into smooth surfaces.

Bending rules

The unexpected smooth shapes observed in experiments and calculated in simulations call for a reconsideration of the composition rules for nanoscale lattice kirigami. The different nanoscale behavior is not because prototypical 2D materials such as graphene, phosphorene, and transition metal dichalcogenides are just more compliant to in-plane stresses. On the length scales of those observations, the materials' υKs are as large as or larger than those of a paper sheet, an assessment that is robust to theory refinements to account for changes in the elastic moduli from quenched disorder or finite temperature.[6] For instance, the graphene υK cited above includes a factor of 10^3 enhancement of its bending moduli as the temperature is increased.

Closer inspection of the graphene landscapes identifies an even deeper conceptual puzzle. Instead of the expected low-strain surfaces separated by stress-focused corners and creases, the smooth surfaces in stiff materials are extremely stress-defocused with low bending energies. That phenomenon is quantified by decomposing the far-field deflection patterns into a sum of separable terms: $f(r,\varphi) = \sum_m f_m(r)e^{im\varphi}$. Those partial wave solutions are individually well described by solutions to the biharmonic equation $\nabla^4 f = 0$, which describes shapes that minimize bending energy. The amplitudes $f_m(r)$ are strongly suppressed for larger angular momentum m, and for low m they are well fit by combinations of two growing far-field radial solutions.

Stress focusing, as expected in elastically stiff media, would produce sharp edges instead and have only a slow, nearly power-law suppression of the large m amplitudes. The directive from lattice kirigami to "bend but don't stretch" is thus replaced by the organic principle to not "fold, spindle, or mutilate," as dictated by the geometrical constraints from the topological defects and external boundaries.

Analyzing the simulations of dislocation-induced 3D shapes reveals three bending rules for kirigami in nanomaterials.[8] First is the law of amplitudes: As mentioned previously, the far-field patterns are well represented by a rapidly converging expansion in the first few cylindrical harmonics m. Second is the law of ratios: Radial solutions to the biharmonic equation contain two terms that grow with distance. For each symmetry-allowed m, the terms always have opposite signs such that the ratio of their amplitudes counterbalances to reach zero area-integrated Gaussian curvature. For example, disks that have been kirigamied, with allowed $m = 2$ solutions, have a shape similar to a potato chip.

Third is the law of boundaries: Although the law of ratios specifies the amplitude ratio, the overall amplitude of the solution is controlled by a boundary energy that's proportional to the perimeter's length. As a result, graphene disks of different radii adopt the same shape so long as the height and lateral position are both scaled to the disk radius. Together the three rules provide a compact, analytic, and reasonably accurate theory of far-field shapes, which can't have a sharp ridge-and-plateau motif.[8]

The bending energy's ascendancy in the problem reveals an oversight in the extrapolation of lattice kirigami to nanomaterials. In continuum theory, Gaussian curvature is a source of in-plane strain, which is unfavorable in any ultrastiff elastic medium. To avoid a stretching energy penalty, which grows faster than the system size, the area integral of the curvature must vanish. But strain couples remote regions of the Gaussian curvature, and with increasing separation, that coupling grows rather than decays; in reciprocal space it diverges[2] as $1/q^4$ as $q \to 0$. That strong coupling at large distances means that the curvature doesn't need to exactly cancel locally. More generally stress-focusing the Gaussian curvature to dislocation/disclination cores is not an inevitable or even a desirable outcome in the shape responses of various nanomaterials. Many 2D materials have taken advantage of that nonlocal physics all along to economize on their bending energies.

Bending rules describe kirigami shapes on finite systems with open boundaries and with the ability to compensate the

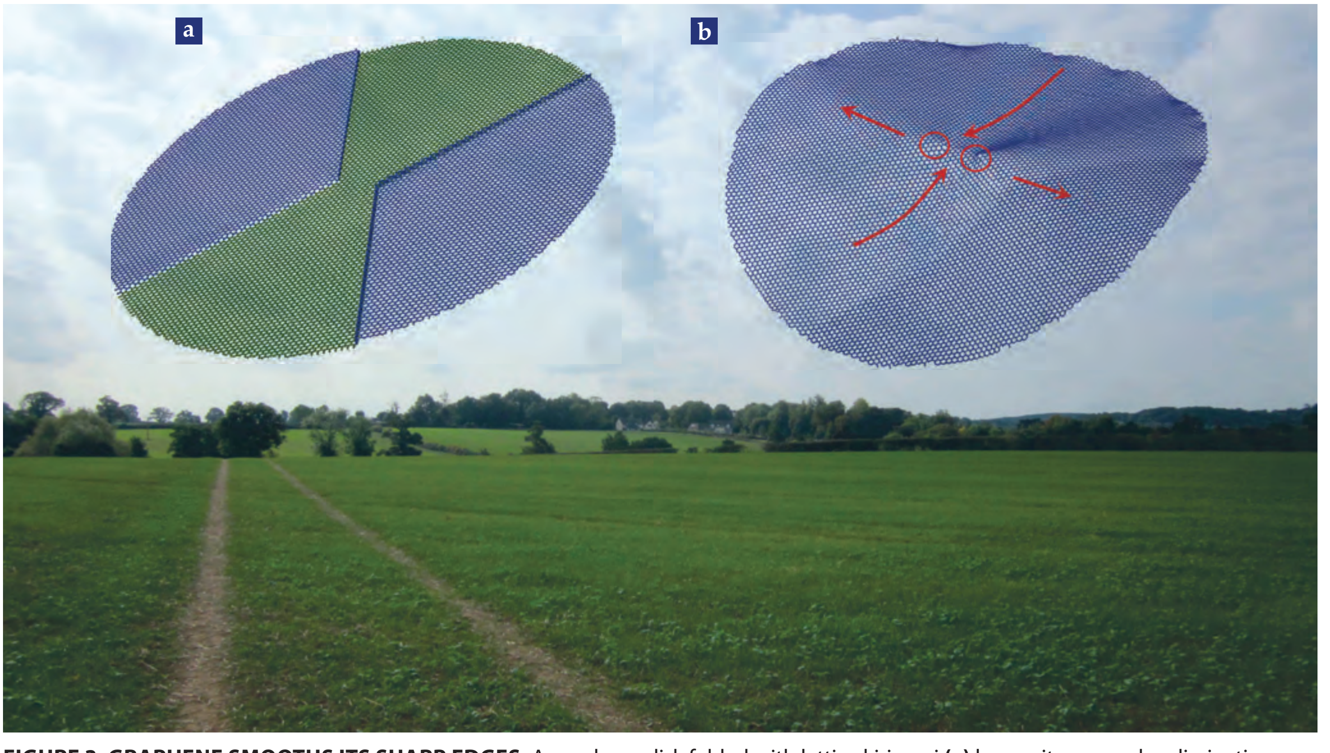


FIGURE 3. GRAPHENE SMOOTHS ITS SHARP EDGES. A graphene disk folded with lattice kirigami **(a)** lowers its energy by eliminating sharp edges. Smooth shapes reduce the bending energy but liberate Gaussian curvature from the corner vertices. The sheet pays a penalty in stretching energy as it strains to achieve a smoothly undulating three-dimensional shape **(b)** evocative of the softly rolling landscape found in the English countryside (background). (Adapted from ref. 8.)

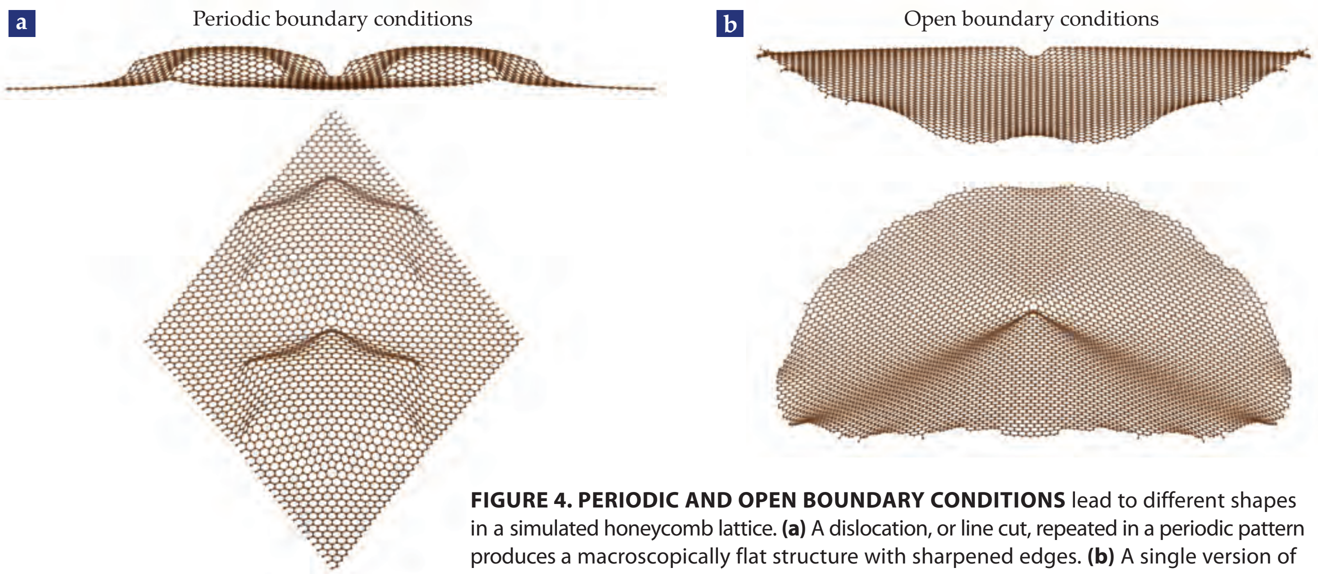


FIGURE 4. PERIODIC AND OPEN BOUNDARY CONDITIONS lead to different shapes in a simulated honeycomb lattice. **(a)** A dislocation, or line cut, repeated in a periodic pattern produces a macroscopically flat structure with sharpened edges. **(b)** A single version of the same cut yields a smoothly warped shape in a disk with open boundary conditions. (Courtesy of Bastien F. Grosso and Eugene J. Mele.)

Gaussian curvature globally. Structures with periodic boundary conditions, which are formed by cutting a periodic pattern of dislocations in an effectively infinite sheet, present an entirely different class of solutions: kirigami without borders. Under periodic boundary conditions with a globally flat surface, the integral of the Gaussian curvature over the simulation area is zero. But the Gaussian curvature field has indirect phonon-mediated interactions, which are present only for the discrete wavelengths that satisfy the periodic boundary conditions, and the $q \to 0$ divergence of the coupling is thereby eliminated.

Atomistic simulations with periodic boundary conditions predict that a nanosheet won't optimize to avoid bending and instead will produce elevated shapes with sharp edges, as shown in figure 4a. Such a graphene microstructure represents a completely new allotrope of "planar" carbon: a defect-stabilized mesophase poised between 2D and 3D. A disk with the same dislocation but open boundary conditions, as in figure 4b, has a smooth 3D surface because the ultralong-range strain coupling returns and the system resumes its primary task of minimizing the bending energy.

Deflection and misdirection

Kirigami can change the electronic behavior of 2D nanomaterials, and graphene is a prototypical example famous for its interesting low-energy electronic behavior (see the article by Andrey Geim and Allan MacDonald, PHYSICS TODAY, August 2007, page 35). That behavior derives from the geometry of two point singularities in its band structure, which occur at two time-reversed momenta, or valleys. At those singularities, known as Dirac points, the bands have linear dispersion relations that form two cones, and the electron and hole bands touch at the shared tip of the cones. Kinematics near the Dirac points follow a solid-state version of the Dirac Hamiltonian for a massless particle.

In graphene that masslessness is enforced by the presence of both time reversal symmetry and twofold rotation symmetry. Uniform in-plane strain preserves both symmetries and leaves the Dirac points intact, although their location in momentum space shifts proportionally to the strain. Microscopically the strain is coupled to the Dirac kinetic energy through a gauge potential that has opposite signs in the two valleys.[11] For a uniformly strained sheet, that vector potential is a pure gauge, and it has no observable consequence. In that case, a strained graphene sheet is still just a sheet of graphene.

A nonuniform strain field can eliminate the band degeneracy, and that lifted degeneracy restores band mass to the electrons and possibly even renders graphene electrically insulating. The curl of the strain-induced vector potential acts like a pseudomagnetic field that pierces closed electron trajectories on the lattice. The strain produced by Gaussian curvature is nonuniform and does restore band mass, although with a catch.[12] In graphene the curl is carried out on a discrete lattice, which endows the Dirac electrons with a mass proportional to the third spatial derivative of the curvature. As a result, a smoothly varying surface-height profile generates a spatially varying mass. The Dirac theory admits the possibility of a negative mass, and here the mass oscillates in sign as a function of angle proportional to $\sin(3\varphi)$. Because of those oscillations, the effect averages out to zero and consequently eludes most measurements made on large length scales.

By inverting that thinking, one can imagine structures that would eliminate the angular oscillations and support instead a spatially uniform pseudomagnetic field. The recipe for doing so is a theorist's dream and an experimentalist's nightmare: The shape's strain field must also oscillate with the same threefold angular anisotropy.[11] In such a system, the lattice curl "eats" its sign changes, and the combination restores a robust nonvanishing contribution in the pseudofield. In practice, the recipe requires alternately pulling and pushing on a graphene sheet in the threefold pattern shown in figure 5a. Electrons propagating on that warped background deflect into orbits as though they were moving in a strong perpendicular magnetic field, although the senses of their orbital circulations are opposite in the two valleys. And since the mechanism for the deflection involves site-to-site hopping and not a conventional magnetic deflection, the magnitude of the equivalent strain-induced pseudomagnetic field can be impressively large, even as high as a few megagauss.

The phenomenon shows up unexpectedly in experiments on graphene sheets grown on nanoscale prismatic platinum islands, shown in figure 5b, which imprint anisotropic strain patterns into the graphene.[13] Spatially resolved scanning tunneling spectroscopy reveals a reorganization of the continuous electronic spectrum of unstrained graphene into resonances corresponding to the quantized Landau levels in the relativistic Hall effect for a ~300 T effective pseudofield.

Graphene kirigami can refine that proof-of-concept demonstration by producing structures that control the size, shape, and symmetry of the strain fields. For example, a threefold symmetric pattern of short dislocation segments (see figure 5c) automatically buckles the sheet into a domed tricorner surface with the requisite spatial symmetry. The deformation releases strain into the dome such that the magnitude and range of the strain-induced pseudofield can be selected by choosing the lengths and orientations of the dislocation segments. Calculations predict related band-structure engineering in other insulating 2D materials, in which strain coupling produces even larger changes to an existing energy gap.[10] The strain could tune the insulators' optical responses or confine the motion of free carriers.

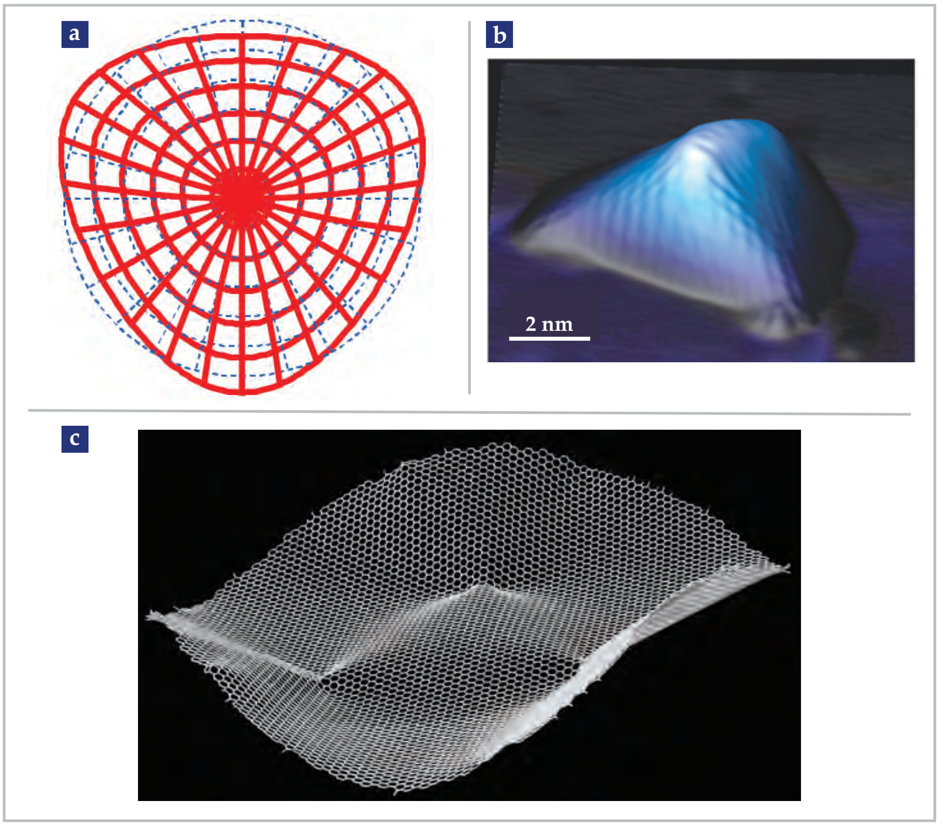


FIGURE 5. STRAIN FIELDS MIMIC MAGNETIC FIELDS. **(a)** A threefold symmetric strain pattern (red) in a graphene sheet produces a smoothly varying pseudomagnetic field, which deflects the electrons like a magnetic field. (Adapted from ref. 11.) **(b)** A strain field with that symmetry occurs in graphene grown on a prismatic platinum island. (Adapted from ref. 13.) **(c)** Graphene kirigami imposes the strain pattern from **(a)** in the tricorner shape resulting from three dislocation pairs at the cusps. (Courtesy of Bastien F. Grosso and Eugene J. Mele.)

Implementing those ideas requires practical methods for cutting and joining defect patterns with specified widths, lengths, and orientations in otherwise ordered nanomaterials. That process is perhaps the biggest distinction between the macroscopic and microscopic forms of kirigami: A literal interpretation of the cut-and-fold instructions is untenable for nanomaterials. Fortunately, some promising alternatives are being developed in other materials contexts. For example, graphene nanoribbons grow from organic precursors with atomically precise control of the ribbon axis.[14] If they're grown on a 2D structure designed to frustrate perfect alignment, topological defects can form in the lattice during the growth process and serve as a scaffold for building a kirigami network.

A related strategy is to directly grow a material on a target 3D shape that has Gaussian curvature. The technique already works for caps of single-layer transition-metal dichalcogenides grown on spherical surfaces.[15] Alternatively, patterning graphene or other 2D materials with functional units can produce hybrid materials that change their shapes in response to external mechanical or chemical stimuli.[16] The strain distributions can, by design, buckle the structure and modify electronic behavior, or, reciprocally, an electronic or optical trigger can produce a shape change *in situ*. Versions of the concept have appeared in flexible electronics made from graphene and silicon.[16,17]

Progress in nano-kirigami is now informed by a confluence of ideas from mathematics, physics, chemistry, materials science, and the arts. That eclectic combination seems poised to turn up yet more cutting-edge science.

We thank Randy Kamien and Marc Miskin for discussions on this topic. Eugene Mele's work is supported by the US Department of Energy under grant no. DE-FG02-84ER45118.

REFERENCES

1. T. Castle et al., *Phys. Rev. Lett.* **113**, 245502 (2014).
2. D. R. Nelson, L. Peliti, *J. Phys. (France)* **48**, 1085 (1987).
3. H. S. Seung, D. R. Nelson, *Phys. Rev. A* **38**, 1005 (1988).
4. T. A. Witten, *Rev. Mod. Phys.* **79**, 643 (2007).
5. M. K. Blees et al., *Nature* **524**, 204 (2015).
6. M. J. Bowick et al., *Phys. Rev. B* **95**, 104109 (2017).
7. O. Lehtinen et al., *Nat. Commun.* **4**, 2098 (2013).
8. B. F. Grosso, E. J. Mele, *Phys. Rev. Lett.* **115**, 195501 (2015).
9. S. Plimpton, *J. Comput. Phys.* **117**, 1 (1995).
10. M. Mehboudi et al., *Proc. Natl. Acad. Sci. USA* **112**, 5888 (2015).
11. F. Guinea, M. I. Katsnelson, A. K. Geim, *Nat. Phys.* **6**, 30 (2010).
12. T. O. Wehling et al., *Europhys. Lett.* **84**, 17003 (2008).
13. N. Levy et al., *Science* **329**, 544 (2010).
14. J. Cai et al., *Nature* **466**, 470 (2010).
15. J. Berry et al., *Nat. Commun.* **10**, 5210 (2019).
16. M. Z. Miskin et al., *Proc. Natl. Acad. Sci USA* **115**, 466 (2018).
17. D.-H. Kim et al., *Science* **320**, 507 (2008).

PT

ISTOCK.COM/4X-IMAGE

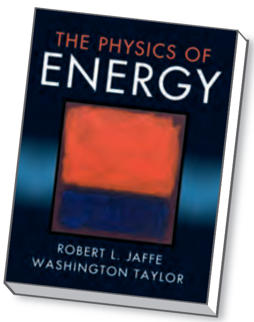
Energy governed by the laws of physics

Robert Jaffe and Washington Taylor's long-needed guide to the fundamentals of energy conversion contains an abundance of knowledge. Energy is an overwhelmingly sophisticated subject. Describing it in both depth and breadth is a Herculean task. The majority of introductory textbooks therefore offer generic, high-level overviews. But the devil is in the details; nuances are important to understanding the big picture and often dictate the feasibility of energy technologies.

As an instructor for an undergraduate course on sustainable energy in a school of engineering and applied sciences, I have had to rely on many texts to cover the subject. I have long wished for a single textbook that could explain energy conversion as it occurs in nature and is utilized in our industrial and digital world in a comprehensive, accurate, and engaging way. *The Physics of Energy* is that book. It covers its subject matter with depth, breadth, care for precision, and clarity.

Jaffe and Taylor have written a textbook made for learning. *The Physics of Energy* emerged from the authors' one-semester course at MIT, which they have taught for more than a decade. Although the book is designed for classroom use, its audience is not limited to engineering and physics students. Sustainable energy sources are receiving enormous societal interest because of the high rate of current energy use and exponentially growing demands, society's heavy dependence on exhaustible fossil fuels, and increased concerns about climate change and the safety of nuclear power. Anyone with an interest in energy will find this book enlightening and informative. It would make a great desk reference for policymakers, as understanding the physical limits of energy use ought to precede any attempt to formulate new energy policies.

The Physics of Energy
Robert L. Jaffe and Washington Taylor
Cambridge U. Press, 2018. $81.99



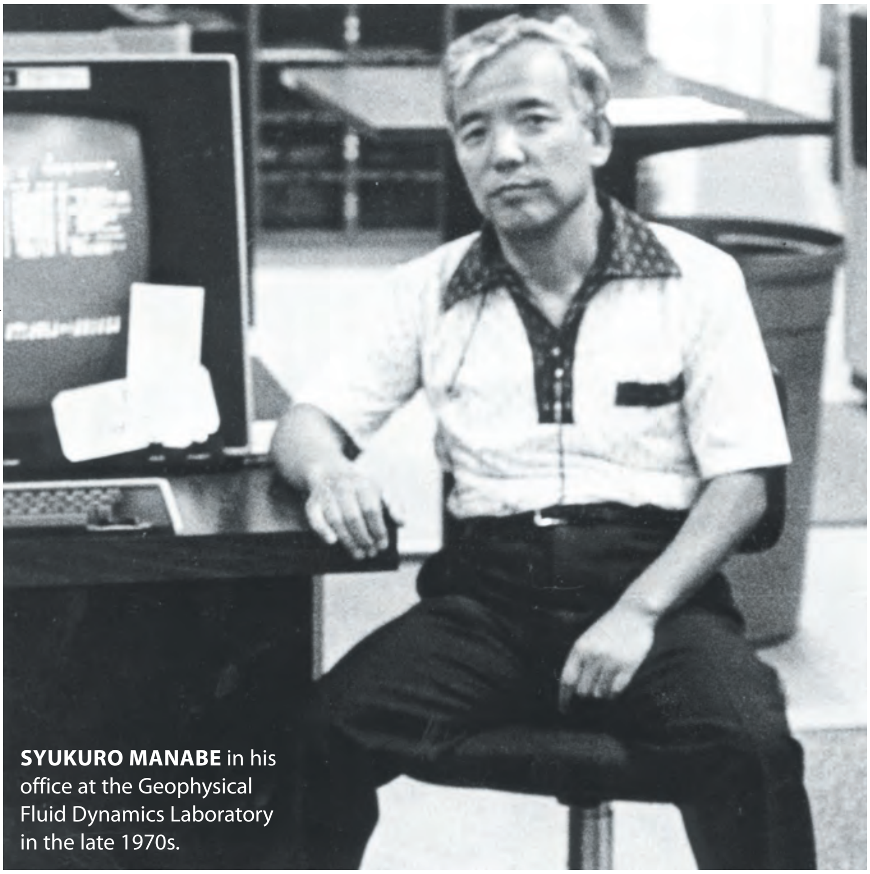
The book contains three logically evolving parts, focused on basic energy physics, sources of energy, and energy system issues and externalities. Energy conversion processes and the physical mechanisms behind them are covered entirely and without redundancies in 874 pages. Jaffe and Taylor artfully connect theory with real-life examples. Have you ever thought of why electric power lines are put underground in most places in Europe but run overhead in North America? Are there any physical limits behind those differences in electric distribution systems? If you are curious, the answer is on page 820. Do you know how important the phase transitions between states of matter are for the functionality of a steam turbine? Chapter 12 explains the physics of phase-change energy conversion and its application to practical devices. The consistent attention to detail is a particularly beautiful aspect of the book.

One of Jaffe and Taylor's greatest contributions is in bridging physics and engineering. *The Physics of Energy* introduces readers to cosmological dark energy, unveils the history of energy in the universe, and covers the most important aspects of sustainable energy technologies. Most textbooks on the subject neglect or deliberately avoid discussing the fundamental meaning of energy and usually rely on a shaky, circular definition of energy as an ability to do work. Those definitions often remind me of Richard Feynman's remark in *Lectures on Physics* that "in physics today, we have no knowledge of what energy is." In contrast, chapter 21 of *The Physics of Energy* includes a comprehensive discussion on the physical meaning of energy and thus gives readers ground beneath their feet in an ocean of energy-related information.

The book also touches on energy efficiency, storage, and challenges of sustainable development. In that context, the one topic that deserved a brighter spotlight is the present and future challenges of solar and wind energy technologies. We will have to contend with the high variability of those energy sources, their inability to provide baseload power, and their low areal power density. All those concerns are especially important in view of rapid urbanization and increased electricity demand for computing, electronic controls, monitoring, artificial intelligence, and other uses of electric power.

To quote Feynman once again, "it is up to the physicist to figure out how to liberate us from the need for having energy. It can be done." Anyone looking for practical recipes on how to create that freedom should keep *The Physics of Energy* at hand.

Julia Mikhailova
Princeton University
Princeton, New Jersey

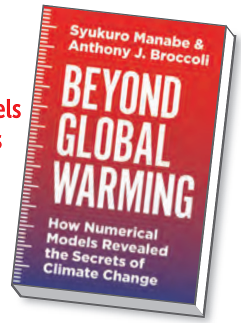


SYUKURO MANABE, COURTESY AIP EMILIO SEGRÈ VISUAL ARCHIVES

SYUKURO MANABE in his office at the Geophysical Fluid Dynamics Laboratory in the late 1970s.

Modeling the future of Earth's climate

Beyond Global Warming
How Numerical Models Revealed the Secrets of Climate Change
Syukuro Manabe and Anthony J. Broccoli
Princeton U. Press, 2020. $35.00



Syukuro Manabe arrived in Washington, DC, in the fall of 1958. He had just finished a PhD at the University of Tokyo and had been invited to work on climate models at the US Weather Bureau. He worked there until 1963, when he moved to Princeton's newly founded Geophysical Fluid Dynamics Laboratory (GFDL), where he spent the rest of his career. He and his colleagues built the first climate model in the US, and Manabe is now regarded as the leading figure in climate model development in this country. Anthony Broccoli worked with Manabe in the 1980s. Together they applied the climate model to various paleoclimate epochs, such as the Last Glacial Maximum. Broccoli then moved to Rutgers University, where he is a professor of environmental sciences. Their new book, *Beyond Global Warming: How Numerical Models Revealed the Secrets of Climate Change*, is a summary of the many major developments that Manabe pioneered, and it stands as a tribute to his career.

Beyond Global Warming begins with the history and basic science behind Manabe's work. The first chapters introduce us to the greenhouse effect and global warming, along with early studies of warming such as the pioneering work of Sweden's Svante Arrhenius in the late 19th century. Arrhenius was the first person to estimate the surface-temperature change resulting from a change in atmospheric carbon dioxide concentration. From there, the book moves on to focus on Manabe's career. It describes his work on one-dimensional models of radiative–convective equilibrium and the early development of global atmosphere general circulation models.

In chapter 5, Manabe and Broccoli describe Manabe's work using early GFDL atmosphere models, including the first computational experiment in which the atmospheric concentration of carbon dioxide was doubled; that experiment is still a standard for all today's climate models. Manabe was one of the first people to use a realistic distribution of land masses and a mixed-layer model of the very upper ocean in his modeling; those firsts strongly influence the spatial pattern of the surface temperature response to carbon dioxide changes.

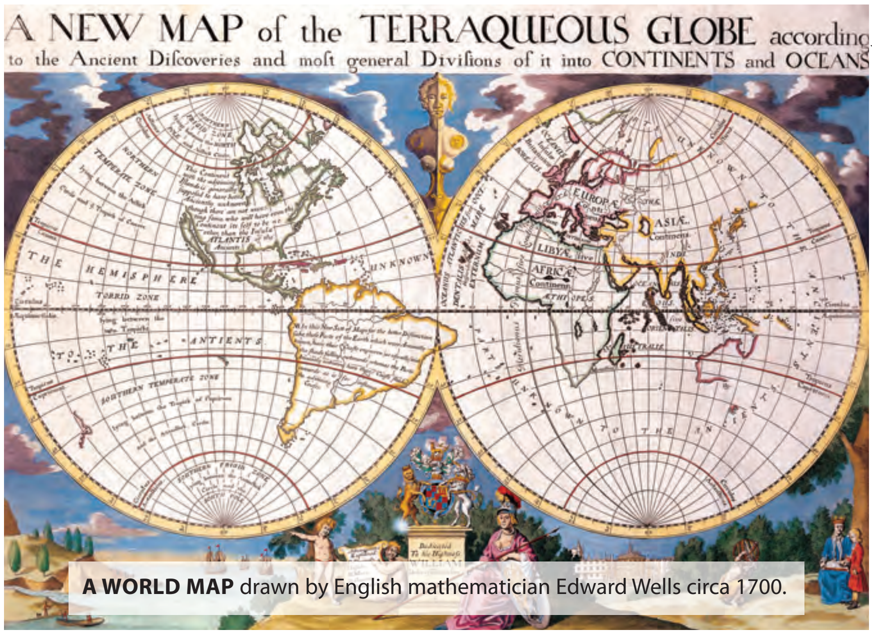
The middle chapters focus on climate sensitivity, which climate scientists define as the equilibrium increase in globally averaged surface temperature due to a doubling of carbon dioxide. Chapter 6 is a general discussion of the various factors that affect climate sensitivity. Manabe was one of the first scientists to show that how clouds were parameterized had a large effect on a model's climate sensitivity. That remains true in all current state-of-the-art climate models. Chapter 7 covers Manabe and Broccoli's work in the mid 1980s, when they determined that the GFDL model had a climate sensitivity of 3.2 °C. In *Beyond Global Warming*, the authors speculate that that number is close to the actual sensitivity of Earth's climate.

The final three chapters highlight the role of the ocean in climate change. Chapter 8 discusses Manabe's work with coupled models using the ocean general circulation model Kirk Bryan developed at GFDL during the 1970s. Manabe and his colleagues suggested early on that the strength of the robust meridional overturning circulation in the North Atlantic Ocean would weaken as the carbon dioxide level increased. That conclusion remains a feature of all future projections made by climate models. Chapter 9 speculates about changes to deep water formation in the ocean when the climate is cold, as at the Last Glacial Maximum. Finally, chapter 10 discusses how the water cycle between the atmosphere and ocean accelerates with global warming as evaporation increases due to higher ocean surface temperatures. As more evaporation leads to increased rainfall, Manabe and Broccoli suggest that wet regions will get wetter and dry regions will get drier.

I have only one point of disagreement with the book, which centers on the discussion of flux adjustments in chapter 8. Flux adjustments are arbitrary changes to the fluxes of heat and fresh water between the atmospheric and oceanic components of a climate model. In early models the adjustments were necessary for a model to maintain a modern climate state as it was integrated forward in time. However, improved climate models, including several developed at GFDL, have been able to run without them for over 20 years now. The need for them indicates inadequacies in the model, which include low spatial resolution, poor parameterizations of important processes, and missing processes. In my opinion, it is preferable to run climate models without using arbitrary flux adjustments, but Manabe and Broccoli endorse their continued use.

This book was a pleasure to read. It is also unlike any other book on climate that I have read because Manabe is unique among climate model developers, and because no other climate book documents a single career. *Beyond Global Warming* will be essential reading for graduate students and postdocs wanting to learn about climate modeling. It will also be an interesting read for any physicist interested in the climate and how future projections are made. It is a fitting tribute to Manabe's career, and I recommend it very highly.

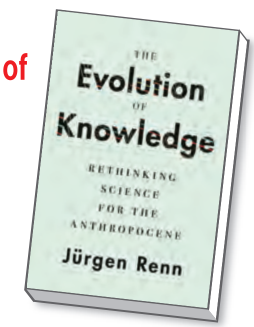
Peter R. Gent
National Center for Atmospheric Research
Boulder, Colorado



CC-PD, VIA WIKIMEDIA COMMONS

A WORLD MAP drawn by English mathematician Edward Wells circa 1700.

A global history of changing knowledge



The Evolution of Knowledge
Rethinking Science for the Anthropocene
Jürgen Renn
Princeton U. Press, 2020. $35.00

Jürgen Renn's *The Evolution of Knowledge: Rethinking Science for the Anthropocene* is a book to read twice, or not at all. A global history of knowledge is a breathtakingly ambitious project, even more so when oriented toward our current global challenges. Renn argues that an expansive perspective is prerequisite for addressing those challenges. His sweeping account of knowledge throughout human history aims to show how the structure of knowledge today contributes to—and offers a platform for addressing—the potentially existential threats of the Anthropocene. Renn faces down the difficulties of crafting such an account with skill and resolve. The result is provocative and challenging and asks much of its readers.

The book's five sections are arranged according to scale, from individuals and communities to institutions to global networks of knowledge. In the book's synoptic first section, "What Is Science?

What Is Knowledge?," Renn argues that knowledge is created by mental, material, and social factors. Our minds determine the categories we can easily apprehend and incline us to certain patterns of abstract thinking. We use mental abstractions to generate representations of ideas, which facilitate both the ideas' manipulation and their dissemination. Human communities are then structured in ways guided by those representations. The interdependence of Renn's three dimensions of knowledge—the mental, the material, and the social—is critical for understanding the book's broader argument and vision.

In part 2, Renn turns to how knowledge changes at the scale of individuals and their local social environments, drawing on developmental psychology and biology. Key here is the insight that abstract reasoning is not a straightforward consequence of our mental architecture; it is shaped by the material and social resources available during cognitive development. That set of resources has distinctive "challenging objects"—phenomena that behave recalcitrantly with respect to contemporary knowledge systems. It also has "borderline problems"—questions that straddle knowledge systems and prompt their extension. Addressing challenging objects and borderline problems is the principal mechanism by which knowledge systems change. The history of knowledge is therefore, according to Renn, contingent, path dependent, and layered, but it proceeds via discernable patterns.

As should now be evident, the book's argument unfolds in a technical language that becomes tractable only with some effort. For example, the book's 18-page glossary (which is unhelpfully broken into thematic sections rather than presented as a single, alphabetized list) defines *scientific knowledge* as "knowledge resulting from the exploration of the potentials inherent in the material or symbolic culture of a society within a knowledge economy specifically dedicated to the generation of such knowledge, allowing for its corrigibility and involving appropriate control procedures." That definition begins to make sense somewhere in part 3, but the understanding is hard won.

Part 3 describes and exemplifies "knowledge economies." Those are the distributed, and sometimes global, networks arranged to preserve, distribute, and employ external representations of knowledge, examples of which range from tools and artifacts to rituals, music, and language. Institutions—which Renn understands broadly as social arrangements, like universities or scientific societies, that reproduce stable forms of behavior within them—are integral components of knowledge economies. They have assumptions and values built into them that both guide and constrain the evolution.

Although Renn's observations about how knowledge moves at institutional scales are edifying, his metaphor of the knowledge economy is problematic. As Renn himself observes, the cultural resources at our disposal shape how we craft the external representations of our abstract concepts. The cultural resources of our age are increasingly those of the market, and that places troublesome constraints on our thinking. Renn, in his conclusion, worries about "new ways of accessing scientific information [being] blocked by its transformation into a commodity." It would therefore be preferable to have a representation of institution-level knowledge processes that did not invite us to think of knowledge as a measurable economic resource to be exchanged and hoarded as capital.

Knowledge on the global scale is the subject of part 4. We are now accustomed to hearing about global processes, but Renn is careful not to miscast global knowledge as distinctive of our time. Our globalized knowledge practices are not a novel legacy of modernity, but part of the layered history of knowledge going back millennia. The point is critical because that history informs the features of our global knowledge system, including the constraints that produce some of our era's challenges.

Those challenges, and how to address them, are the subject of the book's part 5. Renn shows tremendous faith in the power of concepts to do useful work. He argues that there are problems with "the Anthropocene" as a concept, but defends its potential to help us "integrate knowledge from all the disciplines concerned" and address human disruption of the systems stabilizing the biosphere. Doing so, per Renn, requires unshackling ourselves from the constraints of legacy knowledge systems, such as disciplinary divisions, which inhibit the goal-oriented coordination of diverse bodies of knowledge. His proposed recategorization of knowledge into the areas of system, transformation, and orientation knowledge—roughly, the understandings of Earth's natural systems, of human processes and their interaction with those systems, and of human values—is presented with the goal of overcoming intellectual territorialism and fostering the coordinated mobilization of knowledge toward well-defined ends.

That's an abstract proposal, indicative of abstract exposition. Although *The Evolution of Knowledge* is littered with historical examples, the breadth of the subject matter means those examples are necessarily cursory. As a result, it is often challenging to envision how the processes Renn discusses manifest through human agency and, correspondingly, how the solutions proffered can translate into concrete policy. Nevertheless, this book presents a powerful system within which to reason not just about the history of knowledge but about its future. And that is reason enough to read it twice.

Joseph D. Martin
Durham University
Durham, UK

Radioactive

Marjane Satrapi
Amazon Studios/StudioCanal/Working Title Films/Shoebox Films, 2019

In this film based on the life of Marie Curie, the two-time Nobel laureate reflects on her scientific work, her marriage to Pierre Curie, and the scandal surrounding her relationship with fellow scientist Paul Langevin. The film's explanation of the early science of radioactivity is skillfully done—a sequence portraying Marie and Pierre's analysis of pitchblende shows the backbreaking labor of crushing the ore, for instance, and scenes exploring Pierre's scientific interest in spiritualism are evocative and interesting. The confusing time jumps showing events such as cancer therapies in 1957, the Curies as children, and the disaster at Chernobyl, however, are less effective. The movie also makes the strange choice to portray Marie as dismissive of her own accomplishments. *Radioactive* is now streaming on Amazon Prime Video. —MB

NETFLIX

Connected

Latif Nasser, host
Netflix, 2020 (Season 1)

Science journalist Latif Nasser hosts this thought-provoking series about unexpected relationships between people, nature, and technology. The first episode, for example, opens with Nasser watching his young toddler on a baby monitor, then moves to a field biology work site where scientists are using GPS surveillance technologies to understand bird migration. The episode goes on to explore hurricane forecasting, dating app data sets, and the uses of facial recognition for pig farming. Nasser is a charming, high-energy host, and although the show covers a lot of ground, the links between the different topics are clear and unforced. The first season is now streaming on Netflix. —MB

Science Diction

Johanna Mayer
WNYC Studios, 2020 (Season 1)

"Etymology with a side of science" is the catch phrase for this podcast from WNYC Studios and the producers of the weekly public-radio talk show *Science Friday*. Hosted by Johanna Mayer, a digital producer at *Science Friday*, the program reveals the etymology and science connections of certain words and phrases. The first season—which aired during the coronavirus pandemic—featured, not surprisingly, vaccine, quarantine, and Spanish flu. The episodes last about 10–15 minutes and can be accessed on Apple Podcasts, Spotify, and Stitcher. —CC PT

Focus on test, measurement, and analytical equipment

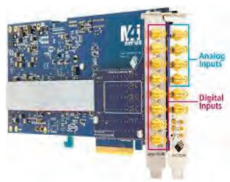
The descriptions of the new products listed in this section are based on information supplied to us by the manufacturers. PHYSICS TODAY can assume no responsibility for their accuracy. For more information about a particular product, visit the website at the end of its description. Please send all new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

Andreas Mandelis

High-performance network analyzers

Keysight Technologies has enhanced its PNA and PNA-X network analyzers to deliver greater flexibility and accuracy while simplifying and speeding complex measurements. The analyzers contain a proprietary low-spurious direct digital synthesis (DDS) source that lets users take accurate measurements with simpler setups and less phase noise interference. With the clean source signals, users can perform two-tone intermodulation distortion measurements with close tone spacing previously only possible with high-performance analog signal generators. The new DDS source also enhances the performance of a wide range of software applications, including modulation distortion and in-phase and quadrature converter measurements, for faster mixer and frequency converter characterization. ***Keysight Technologies Inc**, 1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*

Fast PCIe digitizer input option



Spectrum Instrumentation has announced a new digital input option—the M4i.44xx-DigSMA—that lets its high-speed, high-resolution PCIe digitizers acquire analog and digital signals simultaneously. The option is available for all digitizer models of the company's "44" series with up to 500 MS/s and two or four analog inputs. All models have either two or four analog channels; those with sampling rates of 180 or 250 MS/s have 16-bit resolution, and those with higher speeds of 400 or 500 MS/s have 14-bit resolution. The optional digital input module sits beside the existing digitizer card to provide eight additional digital input lines. It shares a common clock and trigger with the base card, so both the digital and analog inputs are synchronized. That makes the combination suitable for a wide variety of mixed-signal testing applications. ***Spectrum Instrumentation Corp**, 401 Hackensack Ave, 4th Fl, Hackensack, NJ 07601, https://spectrum-instrumentation.com*

Arbitrary waveform transceiver

According to Tabor Electronics, its Proteus system is the world's first arbitrary waveform transceiver. Proteus combines a high-speed arbitrary waveform generator (AWG), a high-speed digitizer, and a field-programmable gate array (FPGA) in a single compact instrument. Based on a PXIe platform, the modular, flexible, and economical system offers high performance, various configuration options, and a user-customizable FPGA for application-specific solutions. Advanced options include real-time data streaming and a fast feedback loop for environment-dependent waveform generation. Proteus is offered in several AWG configurations: with two or four channels and sampling rates up to 1.25 GS/s and 2.5 GS/s or with two channels and sampling rates up to 9 GS/s. A 5.4 GS/s, 8 GHz bandwidth, 12-bit digitizer option offers a complete arbitrary waveform transceiver system. ***Tabor Electronics**, 9 Hatasia St, 3688809 Nesher, Israel, www.taborelec.com*

Tip scanning AFM head for large samples

Park Systems has developed an automated tip scanning head (TSH) suitable for the analysis of next-generation LCD, organic LED, and photonic flat-panel displays larger than 300 mm. The Park NX-TSH's gantry-style tip scanner system moves by means of air-bearing stage technology. Attached to the gantry, the TSH uses the x, y, and z scanners to move to the desired point on a sample, which is fixed on a flexible chuck capable of handling large and heavy objects. Scanning up to 100 μm × 100 μm in the x–y direction and 15 μm in the z direction, the tip produces high-resolution images of the sample's roughness, step height, critical dimensions, and sidewall structure. Paired with an optional integrated microprobe station, which contacts the sample surface and provides current into small devices or chips, the Park NX-TSH uses conductive atomic force microscopy to perform electric-defect analysis. ***Park Systems Inc**, 3040 Olcott St, Santa Clara, CA 95054, https://parksystems.com*

Quantum analyzer with full qubit readout

The Zurich Instruments SHFQA quantum analyzer integrates real-time readout for up to 64 frequency-multiplexed superconducting or spin qubits in one instrument. It is the first quantum analyzer to incorporate a full qubit readout setup, according to the company. The SHFQA addresses readout frequencies up to 8.5 GHz with a clean, wideband spectrum and without the need for mixer calibration. The compact instrument comes with two or four readout channels. Each channel analyzes up to 16 qubits with the highest speed and fidelity by means of the advanced sequencer and the low-latency analysis chain with matched filters, multistate discrimination, and result correlation. The SHFQA is controlled by application programming interfaces or by the company's quantum computing software. A single SHFQA helps to reduce the complexity of small qubit setups; a few synchronized instruments make it possible to scale up to systems of 100 qubits and more. ***Zurich Instruments AG***, *Technoparkstrasse 1, 8005 Zürich, Switzerland, www.zhinst.com*

Optical wavelength meter

The latest addition to Bristol Instruments' 338 optical wavelength meter series combines high measurement accuracy and speed to help optimize optical transceiver testing. The low-cost model 338A uses Michelson interferometer–based technology with FFT analysis to measure the wavelength of CW and modulated signals to an accuracy of ±0.3 pm. It offers a fast measurement rate of 25 Hz, which reduces testing times and improves production throughput. 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*

Digital storage oscilloscopes

Tektronix has introduced its TBS2000B series of digital storage oscilloscopes, which extend the performance of its TBS2000 product series to 200 MHz with a 2 GS/s maximum sample rate. Those features combined with a 5-M-point record length allow the TBS2000B to capture and display more signal so users can debug and validate designs faster. The series offers easy-to-use controls, automated measurements, and built-in instructions that make it suitable for university and training environments. A large 9-inch-wide video graphics array display and 15 horizontal divisions—the most in its class, according to the company—provide 50% more signal visibility. A new lower-noise front-end design offers better signal integrity and more accurate measurements. ***Tektronix Inc***, *14150 SW Karl Braun Dr, PO Box 500, Beaverton, OR 97077, www.tek.com*

High-sensitivity Raman spectrometer

The QE Pro-Raman+ 785 nm spectrometer from Ocean Insight provides low limits of detection for trace-level materials identification. According to the company, its users can detect weak, elusive Raman signatures faster than with other comparable instruments. The spectrometer delivers clean, sharp, stable Raman signatures from 150 cm^{-1} to 3000 cm^{-1}. A cooled, back-thinned detector features long integration times and reduced background fluorescence. Its ability to distinguish sharp peaks from weak Raman spectral signatures makes the QE Pro-Raman+ suitable for analyzing chemicals, pharmaceuticals, illegal drugs, pesticides, and organic materials. The spectrometer can be combined with other components—such as Raman lasers, probes, surface-enhanced Raman scattering substrates, and sample holders—to create a modular system. ***Ocean Insight Inc***, *8060 Bryan Dairy Rd, Largo, FL 33777, www.oceaninsight.com* PT

Eleanor Margaret Burbidge

Elegant, wise, fair, knowledgeable, original, and fiercely determined, Eleanor Margaret Burbidge was one of the great observational astronomers of the past century. One of us (Ostriker) had the privilege of studying under her in the early 1960s at Yerkes Observatory in Wisconsin.

Scientifically, Margaret is best recognized for her lead role in the 1957 paper conclusively arguing that "we are all stardust"—specifically, that most of the heavier chemical elements were made in various phases of stellar evolution. Written by Margaret, husband Geoffrey Burbidge, William Fowler, and Fred Hoyle, the paper laid out the various processes, including both slow and rapid neutron capture, by which the familiar chemical elements were built up. Amazingly, the work's essential arguments remain central to our current understanding of nucleosynthesis. However, the article represents but a small part of Margaret's many accomplishments.

In the 1950s and 1960s, Margaret obtained optically based rotation curves for numerous local spiral galaxies. The curves were characteristically flat, the velocity not declining with radius—a result also found at radio frequencies by Morton Roberts and later in the optical regime by Vera Rubin. Oddly, in the theoretical interpretation of those results, Geoffrey fit the curves' outer parts with a Keplerian $r^{-1/2}$ dependence and made the case for dark-matter halos less obvious—but the results were there for all to see.

In the 1960s Margaret was among the leaders who discovered violent events, including biconical, high-velocity outflowing winds, at the centers of galaxies, which indicated the intense energy input now recognized as active galactic nuclei explosions. Several early papers that Margaret wrote, with Geoffrey, on such observational phenomena were key to late-20th-century revolutionary physical discoveries, including dark matter and black holes. In a 1967 single-author article in *Annual Reviews of Astronomy*, she summarized her work on quasi-stellar objects (QSOs) and described the observational properties, found by her and other investigators, that led astronomers to deduce that massive, accreting black holes were centrally involved. On the velocity dispersions seen in groups and clusters of galaxies, the Burbidges wrote prescient papers that complemented the work of Fritz Zwicky but did not leap to the conclusion that "dark matter" was involved.

Margaret had a knack for finding the facts that challenged the imagination and the conventional view. She was part of the team that in 1974 discovered the high redshift quasar OQ 172, which for many years set the record for the earliest and most distant QSO. It posed a key and still unsolved puzzle over the physical mechanisms that could have operated at that energy scale and early epoch.

Margaret was born in Davenport, UK, on 12 August 1919. In primary school, she was fascinated with large numbers. As a teenager, she read books by James Jeans, a distant relative, and realized that the distances of the stars were large, like the numbers she relished.

At high school in London, Margaret studied physics and mathematics and decided in 1936 to major in astronomy at University College London (UCL). She studied spectroscopy with Herbert Dingle at Imperial College and did her graduate work at the University of London Observatory on spectroscopy of Be stars. She met Geoffrey in a UCL class in 1947; they married in 1948 and had a daughter in 1956.

The Burbidges stayed at the observatory until summer 1951, when they expanded their observational horizons at Haute-Provence Observatory in France. Later that year, on a two-year visa, Margaret went to Yerkes Observatory, and Geoffrey went to Harvard University before joining her at Yerkes in 1952. Margaret used the 82-inch telescope at McDonald Observatory in Texas for spectroscopy of B and Ap stars. A 1953 conference at Yerkes on the origin of elements inspired her to study element abundances; after her and Geoffrey's move to Cambridge University, the two collaborated with Hoyle and Fowler on their famous B^2FH paper on stellar nucleosynthesis.

They returned to the US in 1955—Geoffrey to Mount Wilson Observatory and Margaret to Caltech because of Mt Wilson's ban on women observers. They went back to Yerkes in 1957 and collaborated with Kevin Prendergast on the rotation curves of galaxies. That led to their interest in galactic nuclei and quasars. They moved to the University of California, San Diego (UCSD), in 1962 and returned to Cambridge in the summers to work with Hoyle. Margaret continued her research until about 2006.

AIP ESVA/CLAYTON COLLECTION

Eleanor Margaret Burbidge

A leader in raising awareness of gender and other kinds of discrimination in astronomy, Margaret successfully overcame barriers to women's use of major telescopes and, in her leadership roles, campaigned for gender equality. In 1971 she turned down the Annie Jump Cannon Award from the American Astronomical Society because she felt its restriction to women recipients was discriminatory.

Among her many recognitions, Margaret received the National Medal of Science in 1983. In 1973–75, she was the first, and only, woman director of the Royal Greenwich Observatory, and in 1976–78 she was the first woman president of the American Astronomical Society. From 1979 to 1988, Margaret was the first director of the UCSD Center for Astrophysics and Space Sciences. She died in San Francisco on 5 April 2020.

Jeremiah Ostriker
Princeton University
Princeton, New Jersey
Kenneth Freeman
Australian National University
Canberra

George Trilling

George Trilling, a professor emeritus at the University of California, Berkeley, died in Berkeley on 30 April 2020. His career spanned the development of modern particle physics, from cloud chambers to particle colliders.

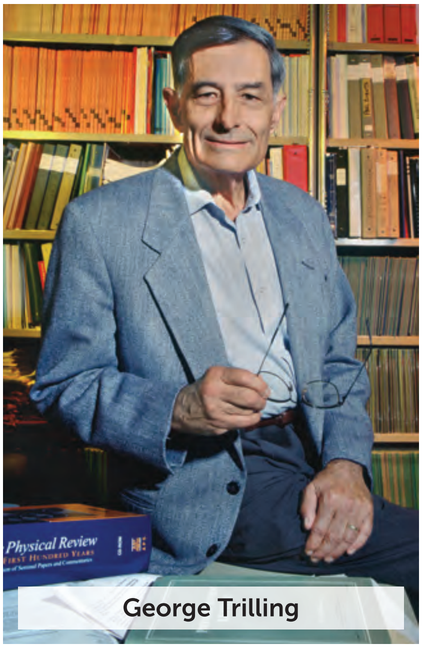
George was born in Białystok, Poland, on 18 September 1930, and his family emigrated to France a few months later. They lived primarily in Nice until 1940, when World War II forced them to emigrate again. The family ultimately settled down near Pasadena, California.

As an undergraduate at Caltech, George became involved in cosmic-ray research and used cloud chambers as part of the group led by Carl Anderson, discoverer of the positron and the muon. After graduating from Caltech in 1951 with a degree in electrical engineering, George continued there under Anderson and studied strange particle decays in a cloud chamber for his thesis, completed in 1955.

After a one-year postdoctoral position at Caltech, George was appointed as an assistant professor at the University of Michigan, but before taking up the post, he spent a year in Paris on a Fulbright fellowship in the group of Louis Leprince-Ringuet. On returning to Ann Arbor in 1957, George worked with Donald Glaser, who had transformed particle physics with his invention of the bubble chamber. It was filled with liquid xenon, ideal for converting photons into electron–positron pairs and thus for observing neutral pi mesons. Glaser's group, which included John Kadyk, John Vander Velde, Daniel Sinclair, and John Brown, made extensive studies of various modes of K-meson decay.

Glaser moved to Berkeley in 1959, and a year later George was offered a tenured associate professorship there in the physics department. Kadyk and Brown joined them to continue with their K-meson decay studies. After Glaser's interest switched to biophysics, George joined with Gerson Goldhaber and Sulamith Goldhaber as a coleader of the Trilling–Goldhaber group, which explored K-meson-induced reactions in bubble chambers at the Lawrence Berkeley Laboratory (LBL) Bevatron, at Brookhaven National Laboratory, and later, after Sulamith's tragic death in 1965, at SLAC.

LAWRENCE BERKELEY NATIONAL LABORATORY



George Trilling

Popular as a teacher, George quickly became a sought-after mentor. His intellectual abilities, judiciousness, and admirable personal qualities led to his appointment as chair of the Berkeley department of physics at age 38. George maintained an active research program while effectively leading the department from 1968 to 1972, through a period of great turmoil associated with student protests on campus.

In 1972 William Chinowsky and the Trilling–Goldhaber group joined Burton Richter's and Martin Perl's groups at SLAC to construct the SLAC–LBL Magnetic Detector for use at the SPEAR e^+e^- colliding-beam facility. Later called Mark I, it was the first detector designed to use full 4π angular detection at colliding-beam facilities. Its design concepts remain the basis for detectors at both e^+e^- and proton–proton colliders. George contributed critically to the success of the Mark I program; he was a wise and effective leader, a brilliant data analyst, and a publication editor extraordinaire. He wrote the analysis code that transformed the tracking-chamber measurements into identified, quantified charged tracks emanating from the collision vertex. Given the novelty of the 4π environment, it was a groundbreaking contribution. It was critical to the discovery of the J/ψ, charm quark, and τ lepton.

The SLAC–LBL collaboration then constructed the Mark II detector, and the group operated it first at SPEAR, then at its higher-energy successor PEP, and after considerable upgrade, at the SLAC Linear Collider. George was cospokesman with Gary Feldman and one of us (Dorfan) for the Mark II upgrade.

From 1984 to 1987, George was the director of the physics division at LBL. He guided the development of semiconductor detectors for colliders, a theme that has continued to the present at LBL.

Following his work at SLAC, George turned to the emerging effort to build the Superconducting Super Collider (SSC) and became the spokesman for the Solenoidal Detector Collaboration (SDC), one of the two experiments planned for the SSC. He guided the SDC's development until the termination of the SSC in October 1993. George retired from active teaching in 1994, but he was instrumental in arranging US participation in the Large Hadron Collider at CERN.

George brought to data analysis exceptional standards of rigor. He was known for his deep understanding of physics and his ability to express it in clear prose. Widely respected as a leader in the physics community, he served as president of the American Physical Society in 2001.

In George, deep intellect and great integrity were coupled with an innate modesty. Always willing to listen, always ready to be there when needed, George had an indelible impact on those who had the great privilege and pleasure of knowing him.

Robert N. Cahn
Lawrence Berkeley National Laboratory
Berkeley, California

Jonathan Dorfan
Stanford University
Stanford, California

Herbert Steiner
University of California, Berkeley

QUICK STUDY

Dave Stevenson is the Marvin L. Goldberger Professor of Planetary Science at the California Institute of Technology in Pasadena.

Juno at Jupiter

David Stevenson

A NASA spacecraft in orbit around Jupiter is yielding new and surprising information about our solar system's dominant planet.

Jupiter is in the class of planets known as gas giants—those planets primarily made from hydrogen and helium gas, which, under gravitational compression, become metallic fluids. It's the most massive planet in our solar system—so massive that it influences everything else, including the delivery of material to Earth. Jupiter is also thought to have formed first among the planets, perhaps a few million years after a cloud of gas and dust from the interstellar medium collapsed and led to our solar system 4.6 billion years ago. That influence means that we need to understand the planet's composition and structure to understand how the solar system came into being.

FIGURE1. THIS VIEW OF JUPITER'S SWIRLING ATMOSPHERE features two storms caught in the act of merging. The two white ovals, left of center in the orange-colored band, are anticyclonic storms that rotate counterclockwise. (Image courtesy of NASA/JPL-Caltech/SwRI/MSSS; image processing by Tanya Oleksuik.)

Through its program of deep space missions, NASA has long recognized the importance of Jupiter. But the planet remains mysterious. Does it have a core? What is the nature of its magnetic field? How do the deep atmosphere and strong winds behave? How much water does it have? To answer those questions, one must visit the planet. For that job, a billion-dollar New Frontiers mission spacecraft named *Juno* was chosen in 2003, launched in 2011, and went into Jovian orbit on 4 July 2016. The mission, led by Scott Bolton of the Southwest Research Institute, is still collecting data. But it's already made some important measurements. This Quick Study outlines three types: The first has to do with gravity; the second, with the magnetic field; and the third, with passive radiation from inside the planet.

Gravity: high harmonics and deep winds

When a spacecraft orbits a planet, its velocity is determined primarily by the gravitational potential it experiences. If Jupiter were spherically symmetric, then *Juno*'s orbit would be explained by Kepler's laws and we would only learn about the planet's mass—assuming we know Newton's constant G from lab experiments, not an easy task. But Jupiter is not spherically symmetric because it rotates, and the rotation is not uniform with latitude and depth.

Jupiter is oblate, a feature that shows up as a gravitational quadrupole and causes a precession in *Juno*'s elliptical orbit. The winds also express themselves in *Juno*'s trajectory. The spacecraft's tracking ability, enabled by a microwave link to NASA's Deep Space Network, is so exquisitely good—better than a factor of 1 part in 10 million in the gravity potential—that scientists can determine higher-order harmonics of the gravity field and discern differences between northern and southern hemispheres caused by the differences in the winds in those regions. The gravity harmonics, together with interior modeling of the planet, have allowed researchers to infer much about Jupiter's internal structure.

The simplest interpretation of the higher harmonics is that the planet has a diluted central concentration of heavy elements rather than a compact central core of them. The difference amounts to a mere 10 to 15 Earth masses, a small fraction of the planet's total mass. Even so, current ideas of Jupiter formation favor the accumulation of a planetary embryo from that tiny mass before the mostly hydrogen envelope accumulated on top. The dilution of those heavy elements and hydrogen is not easily explained. But that interpretation is the first of several surprises *Juno* has provided. One possible explanation is a giant impact during Jupiter's formation, an event that might have stirred up a primordial core.

Before *Juno*, some researchers claimed that the winds in Jupiter's atmosphere are purely meteorological—that is, confined to a thin layer extending a mere 100 km in depth. Others argued they are a surface expression of convective flows that extend as much as 20 000 to 30 000 km deep, limited only by the region where hydrogen gas becomes liquid metal. The truth appears to be somewhere in the middle. Jupiter has strong winds (see figure 1) that extend perhaps 3000 km deep, judging by the harmonics of the gravity field (see PHYSICS TODAY, May 2018, page 19). Below that depth, the winds taper off because of the unavoidable electromotive force that arises from Fara-

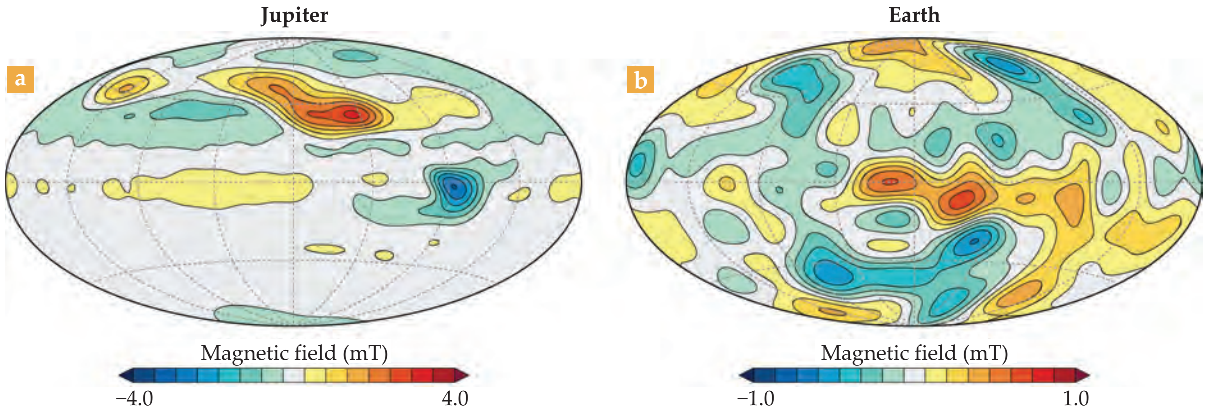


FIGURE 2. JUPITER'S AND EARTH'S MAGNETIC FIELDS. (a) Almost all of Jupiter's nondipole radial field, simulated here at 0.9 of Jupiter's radius, is concentrated in the northern hemisphere. **(b)** Earth's nondipole field, by contrast, is evenly distributed throughout. (Adapted from D. J. Stevenson, *Annu. Rev. Earth Planet. Sci.* **48**, 465, 2020.)

day's law when the electrically conducting wind flows across magnetic field lines.

The consequences of the flows aren't well understood, but we know that for typical winds and field strengths, an electrical conductivity similar to that of salty water or the human body, roughly 1 siemens per meter, leads to a substantial change in field strength. That conductivity is orders of magnitude less than that of a typical metal and occurs at about 97% of Jupiter's radius, well outside the region where the main magnetic field is produced but at a place where the winds appear to decline. Still, planetary scientists lack a dynamical explanation for how the winds decline with depth. The nature of the winds on Jupiter—and those on Saturn, Uranus, Neptune, and exoplanets—will likely remain a lively topic of future research, stimulated by *Juno*'s discoveries.

Magnetic field: northern exposure

Scientists have been aware of Jupiter's large magnetic field since the early days of radio astronomy in the 1950s, and previous spacecraft have documented a field that was somewhat Earthlike (with a similar dipole tilt and similar ratio of quadrupole to dipole) after taking into account a field-generating region that extends closer to the surface than Earth's. However, *Juno* has discovered some unexpected features in the magnetic field, whose spatial distribution is strongly heterogeneous (see figure 2). The planet exhibits sparse but strong fields confined to small regions, most notably the planet's Great Blue Spot—a region of downward flux near the equator.

Learning about magnetic fields in our solar system has been a sobering experience. It seems that whenever we measure anew we find something different, and Jupiter is no exception. Like Earth, however, Jupiter has an internal magnetic field that changes in time. The cause is presumably the advection of magnetic field lines by flow within the planet (see the article by Peter Olson, PHYSICS TODAY, November 2013, page 30). In Jupiter's case, though, the advection could be telling us about the connection between the field and the winds. That connection would suggest an interaction between magnetic and gravity fields.

Microwave radiation: a watery mystery

I've saved the most remarkable discovery for last. It is common practice in the planetary community to think of clouds in a convective atmosphere as being the surface expression of a well-mixed region beneath them, with the base of the cloud deck defined simply by equating the partial pressure of the cloud-forming material to its local vapor pressure. In that picture, ammonia should appear as well mixed within Jupiter deeper than a pressure of about one bar—the surface pressure of Earth.

The microwave radiometer on *Juno* picks up the planet's interior radiation, which is attenuated by intervening ammonia. Astonishingly, though, it did not see that well-mixed state. Instead, it generally saw a factor-of-two depletion in ammonia, except for a narrow band near the equator. If ammonia is not well mixed, what are we to say about water, which, like ammonia, is a condensable species? Water abundance in Jupiter is one of the Juno mission's key goals. Assuming the narrow band is also well mixed in water, Juno's microwave data suggest that water is at least as abundant as it would be if the atmosphere had the same elemental abundances as the Sun. And it might be two or three times more abundant.

Still, we are confronted with the conclusion that what is happening in most of Jupiter's atmosphere is poorly understood. The mystery is confounding for interior models as well because they depend on knowing the atmospheric boundary condition. Lots of ammonia resides at the top of the atmosphere and we expect a lot in the interior, but at issue is how to transport it. The best way to hide a lot of transported ammonia from observation is to deliver it quickly but in a spatially localized way. One proposed solution invokes a form of hail known as mushballs, capable of rapidly delivering ammonia and water to great depths without evaporation. Perhaps future observations can help us unravel that unsettling result.

Juno is ongoing; it has not answered all questions and has raised new ones—a measure of success. Plans are afoot to extend the mission, with more discoveries promised and even a possible trip to the Galilean moons.

Additional resources

▶ S. J. Bolton et al., "Jupiter's interior and deep atmosphere: The initial pole-to-pole passes with the Juno spacecraft," *Science* **356**, 821 (2017).

▶ S. J. Bolton et al., "The Juno mission," *Space Sci. Rev.* **213**, 5 (2017).

▶ D. J. Stevenson, "Jupiter's interior as revealed by Juno," *Annu. Rev. Earth Planet. Sci.* **48**, 465 (2020). PT

BACK SCATTER

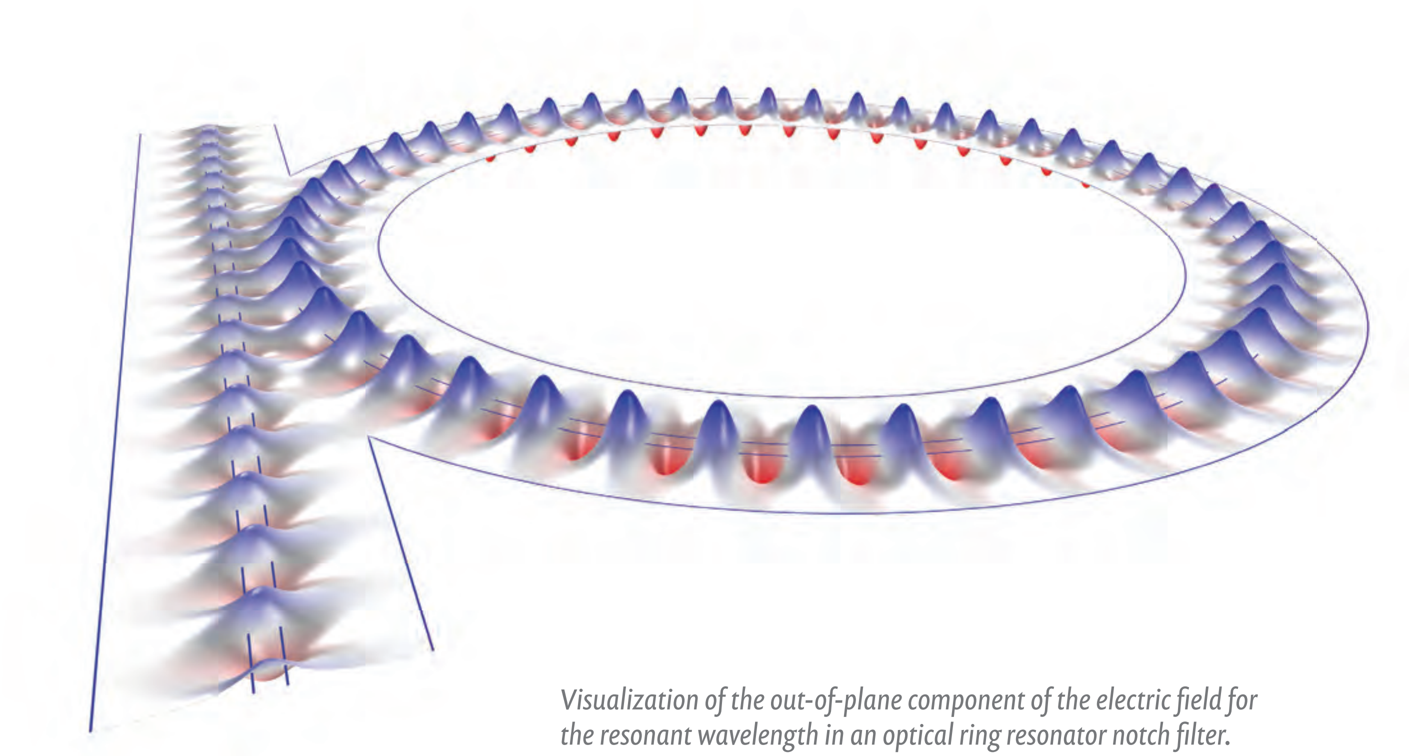
Predicting large solar flares

A solar flare can reach Earth in about eight minutes, and its impulsive radiation can harm astronauts in space and technology on the ground. Physicists have constructed various empirical models to explain the occurrence and properties of solar flares, but their accuracy remains low. Kanya Kusano of Nagoya University in Japan and his colleagues considered solar flares from a mechanistic, physics-based approach. In their concept, magnetic reconnection triggers a double-arc magnetic loop instability that then disperses some of the free energy released during a solar flare. Following that line of reasoning, Kusano and his colleagues developed a flare model that provided lead times of a few hours to an entire day.

One of the flares they used to test their model is pictured here in the upper left area of the image and occurred in 2012. To better understand such large flares, the researchers tuned their model using what they call the κ parameter, which estimates the length scale and the region where the magnetic reconnection can trigger the instability. They also used κ to estimate the magnetic free energy that can be released by the double-arc loop. When the researchers compared their simulations with magnetic location data collected by NASA's *Solar Dynamics Observatory*, they found the model successfully predicted the occurrence and size of seven of the last nine largest solar flares. (K. Kusano et al., *Science* **369**, 587, 2020; image courtesy of NASA's Goddard Space Flight Center/*SDO*.) —AL

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

It all started with two buckets of water...



Visualization of the out-of-plane component of the electric field for the resonant wavelength in an optical ring resonator notch filter.

In 1870, a scientist named John Tyndall tried to control light using two buckets of water, illustrating total internal reflection to a fascinated audience. Today, researchers have more advanced tools at their disposal. When fabricating and analyzing optical waveguide prototypes, modern-day engineers can use numerical simulation software to speed up the design process.

The COMSOL Multiphysics® software is used for simulating designs, devices, and processes in all fields of engineering, manufacturing, and scientific research. See how you can apply it to designing silicon waveguides.

comsol.blog/silicon-photonics

COMSOL