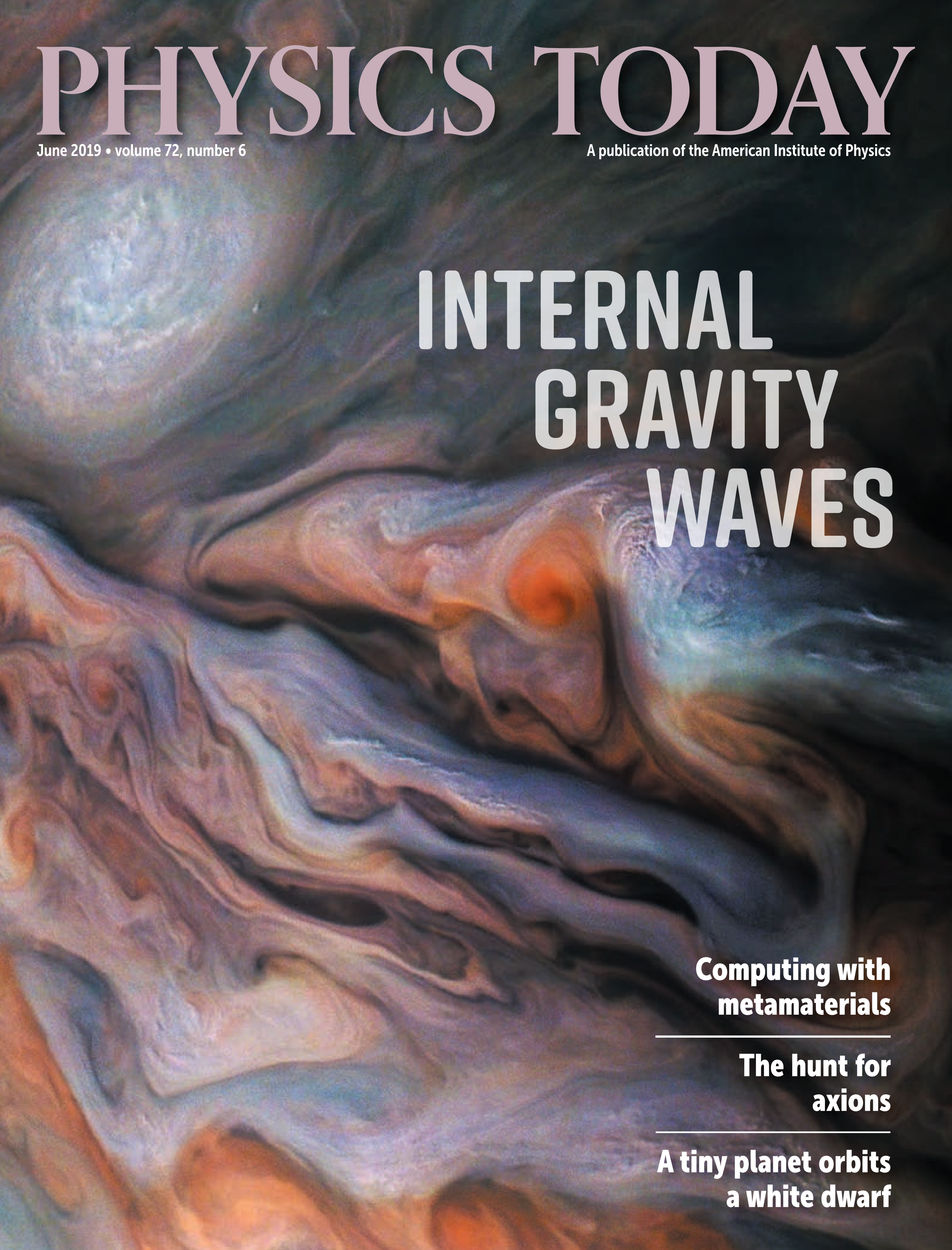
PHYSICS TODAY

June 2019 • volume 72, number 6

A publication of the American Institute of Physics

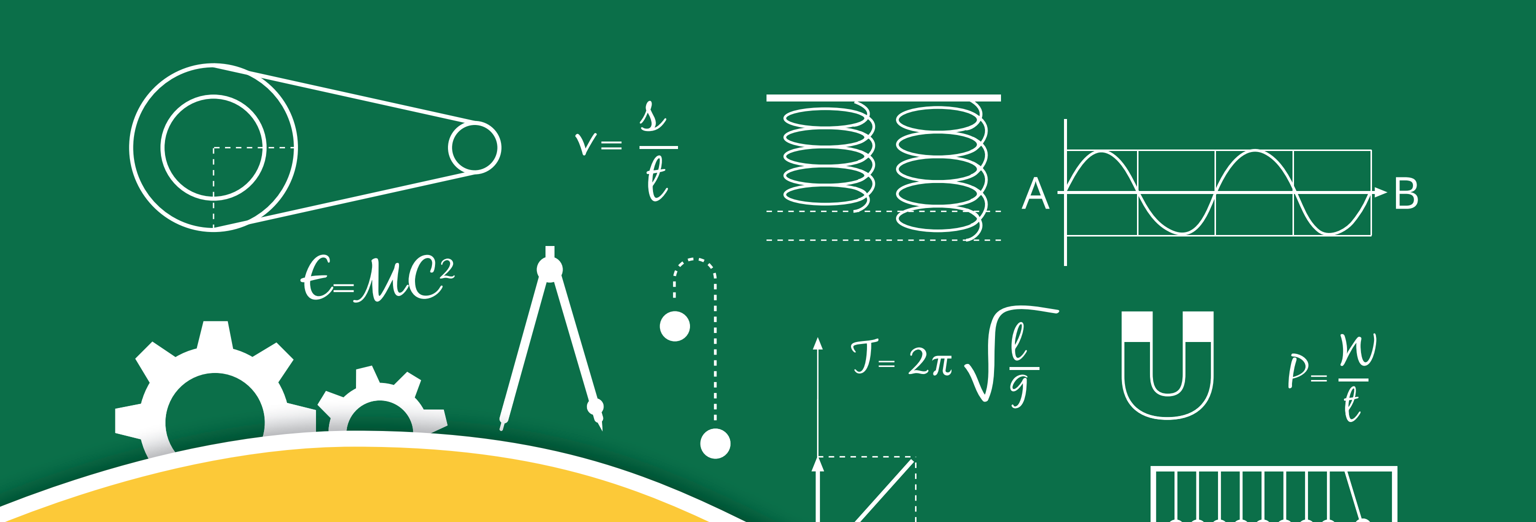


INTERNAL GRAVITY WAVES

Computing with metamaterials

The hunt for axions

A tiny planet orbits a white dwarf



Physics Today
helps physics teachers stay up to date when designing their curriculums.

I use it to keep informed about new discoveries, results and historical background. It is extremely useful for teaching my physics classes.

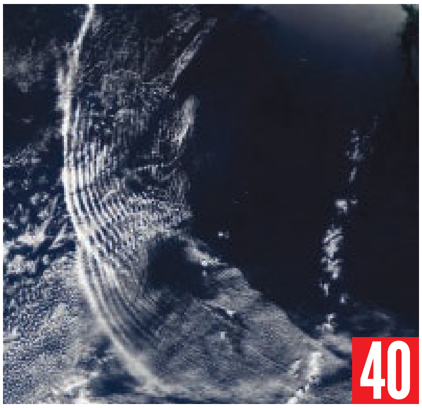
I generally use *Physics Today* to get updates about the recent trends and research breakthroughs, which sometimes helps in teaching Master's students.

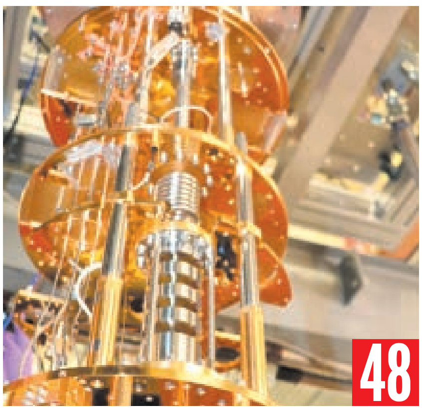
PHYSICS TODAY

Figures and quotes obtained from a *Physics Today* reader survey.

PHYSICS TODAY

June 2019 | volume 72 number 6





FEATURES

SPECIAL FOCUS ON INTERNAL WAVES

34 Spontaneous generation of internal waves

Callum J. Shakespeare

The waves responsible for transporting energy in the ocean and atmosphere are not solely a product of the combined forces of winds, tides, and mountains.

40 Obscure waves in planetary atmospheres

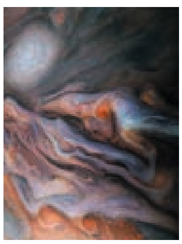
Erdal Yiğit and Alexander S. Medvedev

On Earth and on other planets, internal gravity waves shape the dynamics and thermodynamics of the atmosphere.

48 Putting the squeeze on axions

Karl van Bibber, Konrad Lehnert, and Aaron Chou

Microwave cavity experiments make a quantum leap in the search for the dark matter of the universe.

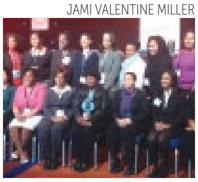


ON THE COVER: Images of Jupiter's swirling atmosphere, such as the one shown here taken by NASA's *Juno* spacecraft, have provided an unprecedented look at the gas giant's atmosphere. On **page 40**, Erdal Yiğit and Alexander Medvedev discuss the role of gravity waves in the atmospheric dynamics of Jupiter, Earth, and other planets. (Image courtesy of NASA/JPL-Caltech/SwRI/MSSS/Gerald Eichstädt/Seán Doran.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org

JAMI VALENTINE MILLER



▶ Black women in physics

Since 1972 more than 60 000 people have received PhDs in physics in the US; fewer than 100 are African American women. One of them, Jami Valentine Miller, writes about the organization she founded to honor the pioneering physics PhD holders, inspire future physicists, and promote diversity in the field.
physicstoday.org/Jun2019a

WHOI

▶ Marine seismology

Among the thousands of seismic stations collecting data globally, the vast majority are on land. That's why some seismologists are developing instruments that listen in on subsurface vibrations from the ocean depths. Physics Today's Alex Lopatka reports on how researchers are filling in the gaps as to what goes on beneath Earth's oceanic crust.
physicstoday.org/Jun2019b

ED BATRES

▶ Popular science

Awesome Con might seem like an odd place to find an astrophysicist. But for Erin Macdonald, fan conventions are a second home. The former LIGO researcher talks to Physics Today about communicating science in myriad venues and advising TV producers and novelists about science.
physicstoday.org/Jun2019c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 72, number 6. Published monthly by the American Institute of Physics, 1305 Walt Whitman Rd, Suite 300, Melville, NY 11747-4300. Periodicals postage paid at Huntington Station, NY, and at additional mailing offices. POSTMASTER: Send address changes to Physics Today, American Institute of Physics, 1305 Walt Whitman Rd, Suite 300, Melville, NY 11747-4300. Views expressed in Physics Today and on its website are those of the authors and not necessarily those of AIP or any of its member societies.

PHYSICS TODAY

www.physicstoday.org

DEPARTMENTS

8 From the editor

10 Readers' forum

Commentary: Improving diversity and inclusion in STEM graduate education — *Alexander L. Rudolph* • Letters

14 Search & discovery

Iron-rich object closely orbits a white dwarf • Inverted kinetics seen in concerted charge transfer • Flows of volcanic rock and gas ride a carpet of air • A metamaterial solves an integral equation

26 Issues & events

Physical and mathematical approaches yield insights into how cancer develops and spreads • Questions surround NASA's shutdown of an international cosmic-ray instrument

56 Books

The hunt for natural quasicrystals — *Michael Engel* • Telling time in Tokugawa Japan — *Sara J. Schechner* • Electromagnetism textbook bridges the gap between basic and advanced — *Savan Kharel* • New books & media

61 New products

Focus on test, measurement, software, and instrumentation

67 Obituaries

Manfred Eigen

70 Quick study

Indestructible plasma optics — *Howard Milchberg*

72 Back scatter

Spore vortex

FROM THE EDITOR

Here come the robot authors

Charles Day

"Lithium-ion batteries (Li-ion batteries) have been commonly used as power sources in consumer electronics including laptops, cellular phones, and full and hybrid electric vehicles because of their long cycling life, high energy capacity, and eco-friendliness. Considerable efforts have been devised to examine useful electrode materials for Li-ion batteries with long cycle life and high capacity."

Thus begins *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research* (2019). The 278-page book is the first from Springer Nature that has a machine, not human, author. Named Beta Writer, the robot author is an algorithm developed by computer scientists and editorial subject-matter experts at Springer Nature and Goethe University Frankfurt in Germany.

In the book's human-written preface, Goethe University's Christian Chiarcos and Niko Schenk reveal that Beta Writer combines two subfields of artificial intelligence: natural language processing and machine learning. The raw material for *Lithium-Ion Batteries* consists of 1086 scientific papers. After processing them to index and standardize their content, Beta Writer set to work identifying common themes. The result was an ordered set of chapters, sections, and subsections. Producing the text for the book involved analyzing, simplifying, and synthesizing text from the papers. For example, each chapter's introduction and conclusion were generated from the introductions and conclusions of the papers included in the chapter.

How successful was the enterprise? Whereas each chapter starts with an introduction, the book itself lacks a machine-written preface. My hunch is that providing an overview of an entire field is beyond Beta Writer's current ability, possibly because writing a good overview entails making strong, idiosyncratic choices about what to include and exclude. Also likely beyond Beta Writer's ability: explaining why it wrote the book in the first place. Not surprisingly, given how the chapters were generated, their titles tend to be lists of technical terms, some of which are puzzling or obscure. The fourth chapter bears the title "Models, SOC, Maximum, Time, Cell, Data, Parameters."

On the other hand, the writing itself, though dull, serves for the most part to achieve one of the book's objectives: To organize those 1086 papers into a coherent, structured, abstracted whole. Indeed, the algorithm's most likely use-case (to use a term from the software industry) is to trawl through literature on a given topic—topological insulators, say—and distill and arrange its main findings. Beta Writer should prove lucrative for Springer Nature.

What should we make of the fact that Beta Writer more or less successfully reviewed a substantial scientific area? When I examined the author guidelines of a handful of journals in physics and its closely related sciences, I discovered that none of them explicitly require papers to have an introduction. Yet all the papers I encountered had one. You can imagine that papers about cathodes made from lithium manganese oxide (one of the subsections in Beta Writer's book) might all begin with similar introductions that cite the same prior papers. And you might also imagine that each paper reports an improvement of some kind in the cathode's performance. Insofar as the paper is a thick wrapper of words around a set of quantitative results, is it any wonder that Beta Writer could successfully extract results and put them in context?

Beta Writer isn't the only robot science writer in town. Two years ago, a company called sciNote launched an AI tool within its electronic notebook software, ELN. Once a scientist has all of his or her data and lab notes in ELN, the Manuscript Writer tool intervenes to write a report. A visit to sciNote's website revealed this quote from a contented user: "Not only does the new feature generate manuscripts quickly, it also provides several versions that can be used to assemble that perfect publication for your data."

The formulaic style of modern scientific papers makes the work of Beta Writer and Manuscript Writer possible. The style is efficient because it's predictable. But if 21st-century scientists want to engage readers in addition to informing them, they might consider adopting the personal, conversational style of their 19th-century predecessors. Lord Rayleigh opened his derivation of what became known as Rayleigh scattering with the following lines:

> It is now, I believe, generally admitted that the light which we receive from the clear sky is due in one way or another to small suspended particles which divert the light from its regular course. On this point the experiments of Tyndall with precipitated clouds seem quite decisive.[1]

How long will it be before robots can write like that?

Reference

1. J. W. Strutt, *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **41**, 107 (1871).

PT

Commentary

Improving diversity and inclusion in STEM graduate education

SCHOLARS FROM CAL-BRIDGE, a joint California State University–University of California bridge program to help physics and astronomy students from participating campuses pursue a PhD in astronomy, physics, or a related field. Although Cal-Bridge targets UC PhD programs, about 60% of the scholars go to PhD programs at other institutions. (Photo courtesy of Kerean Povich.)

Let's say you are a faculty member in a physics or astronomy department and for the first time will be serving on the graduate admissions committee. You have heard a lot about the problems our fields have with diversity. For example, questions have been raised about the usefulness of the GRE exam in determining who should be accepted to PhD programs. Many of your colleagues believe the GRE says "something" about candidates' abilities, and they ask, "How will we get through hundreds of applications without some objective metric to sort them?" Vaguely uneasy with those arguments, you believe something should be done, but you are not sure what.

The uncertainty, discomfort, and downright fear around addressing issues of inclusion constitute real and unnecessary barriers to increasing diversity in our fields. The fields of STEM (science, technology, engineering, and mathematics) have had a glaring diversity problem for decades, and progress has been uneven. According to an NSF report,[1] only 37% of PhDs awarded in STEM in 2016 went to women. In astronomy the figure is even lower, at 31%; and in physics the number is lower still, at 18%.

The situation is much worse for people from underrepresented minority groups, including American Indian or Native Alaskan, Black or African American, and Hispanic or Latinx students. In 2016, 12% of STEM PhDs awarded to US citizens and permanent residents went to members of those groups, and only 6% in physics and astronomy. Those numbers are well below 31%, the US Census Bureau's estimate of those groups' representation in the US population that year.

But back to your dilemma. You need help convincing your colleagues on the admissions committee, and you will find it in the recent report by the American Astronomical Society (AAS) Task Force on Diversity and Inclusion in Astronomy Graduate Education,[2] cochaired by myself and Gibor Basri of the University of California, Berkeley. The report provides a road map for addressing known barriers to diversity and inclusion in astronomy, most of which are also found in physics and other STEM fields. The major barriers include admissions practices, such as the GRE exam, that are based on flawed metrics and weed out nontraditional candidates; low retention rates caused by department climates that are not always welcoming to women and people of color; and the lack of systematic data to track progress.

To address admissions, retention, and data collection, the AAS task force divided into three working groups, one to address each area. In the end, the entire task force met to integrate the work of the three groups into a coherent whole. The report is designed to be a practical guide for departments, with specific recommendations in each area. More than half of the report is appendices containing detailed recommendations and tools that some departments are beginning to use to improve their admissions practices and department climate.

The task force included members of the astronomy community from a wide variety of schools and programs: research universities, minority- and Hispanic-serving institutions, and bridge programs that help students from traditionally underrepresented groups matriculate into PhD programs. In addition, three nationally recognized social-science experts advised the task force, one on each area of concern. Together they made sure that all recommendations were based on evidence or best practices.

As a new member of the admissions committee, you can use the AAS report to become a well-prepared, informed advocate for diversity. Specifically, you can take into committee meetings the report's key recommendations pertaining to admissions, such as "implement evidence-based, systematic, holistic approaches to graduate admissions, based on the existing literature as well as on self-study when possible" (reference 2, page 12).

But what is holistic admissions, and how can it be implemented? The report describes a three-component model that was developed by one of the task force advisers: "Holistic review in graduate admissions should be 1. *comprehensive*, considering a variety of student qualities including their socioemotional/noncognitive competencies, 2. *systematic*, articulating how reviewers should look for these qualities, and 3. *contextualized*, considering how students' characteristics and achievements reflect not only their potential, but also the opportunities they have had, their developmental trajectories, and known sources of error in standard metrics" (reference 2, page 12).

The key idea is that traditional measures used in admissions are incomplete and do not weigh characteristics for success as a physics or astronomy researcher as opposed to as a classroom student. Those characteristics include perseverance, creativity, conscientiousness, realistic self-appraisal, a focus on long-term goals, and leadership.

A powerful new study of various factors contributing to PhD completion looked at more than 2000 US students receiving physics PhDs from 27 programs over a 10-year period.[3] It found that the physics and verbal GRE tests showed no statistically significant relationship with PhD completion. The range of physics scores varied from the 10th percentile to the 90th, so the lack of correlation is not due to a restricted sampling range.

The GRE-Q (quantitative measure) showed a barely statistically significant correlation with PhD completion. Students scoring in the 90th percentile for the GRE-Q are only 9% more likely to receive their degree than those scoring in the 10th percentile, so even that test is a poor tool for predicting success in graduate school. The use of the GRE for PhD admissions becomes even more problematic when one considers that scores on all three GRE tests—physics, verbal, and quantitative—show strong correlations with gender and ethnicity in a way that greatly reduces diversity.[3,4]

So, without the GRE to guide you, how do you make admissions decisions? The AAS report suggests that "programs should reduce reliance on standardized tests, structure information gathered via recommendation letters, and incorporate assessment of socioemotional competencies (i.e., non-cognitive skills). Faculty reviewers should also approach prospective students as learners, not only as research or teaching assistants, and evaluate them for their potential to grow into great scientists, not only for their accomplishments to date. *Because opportunities to learn and conduct research vary considerably with forms of social privilege, it is critical that programs working to mitigate inequalities not simply admit the students with the most impressive credentials*" (reference 2, page 13; emphasis added).

In particular, the use of rubrics to evaluate candidates can ensure that reviewers consider the many characteristics of successful PhD students, including the socioemotional competencies mentioned above. Toolkits, some of which are included in the appendices of the AAS report, can guide admissions committees in assessing these skills as a complement to more traditional measures such as GPA, essays, and letters of recommendation.

The AAS report contains examples from PhD programs whose holistic admissions practices have begun to show success in boosting diversity without reducing student quality. In fact, evidence from some of the programs suggests that the attention to socioemotional skills has increased the quality of PhD students. Although implementing such practices may take more work than simply sorting by physics GRE score, that extra effort should vastly improve the resulting PhD pool.

The time is long past to make the physics and astronomy communities representative of the society we live in, and thereby utilize the full potential of society's scientific ability. Multiple factors have impeded progress in achieving that goal, but the largest is probably ignorance of the real characteristics that influence success in graduate school. Most physics and astronomy faculty members assume they know what a successful PhD student looks like, but such assumptions are largely untrue or untested. If physics and astronomy faculty seek out better information and implement the types of practices recommended in the AAS report, then we may yet succeed in our shared goal of improving diversity and inclusion in our fields.

References

1. National Center for Science and Engineering Statistics, Directorate for Social, Behavioral and Economic Sciences, *2016 Doctorate Recipients from U.S. Universities*, NSF (2018).
2. A. Rudolph et al., *Final Report of the 2018 AAS Task Force on Diversity and Inclusion in Astronomy Graduate Education*, American Astronomical Society (2019).
3. C. W. Miller et al., *Sci. Adv.* **5**, eaat7550 (2019).
4. C. Miller, K. Stassun, *Nature* **510**, 303 (2014).

Alexander L. Rudolph
(alrudolph@cpp.edu)
California State Polytechnic University Pomona

LETTERS

For love and physics

Toni Feder's photo story "Snapshots from the life of Cécile DeWitt-Morette" on the Physics Today website (10 October 2017) brought back fond memories about my mother and more than a few smiles! I would like to correct one anecdote, though, which I gather came from my sister Chris when she provided background on our mother without knowing it was incorrect.

The error, I know, would have saddened Cécile had she read it. It had to do with her reason for not marrying Peng Huan-wu, her adviser at the Dublin Institute for Advanced Studies in 1947. The story says it was because he wasn't French, but that was not the case.

Cécile did indeed tell us she used the excuse that our father, Bryce DeWitt, was not French (or Catholic) for her initial reluctance to marry him, though they wed in 1951. But her reasons for not marrying Peng were quite different. She was very much in love with him and would have married him, but it was the late 1940s and he was returning to China, which was in the midst of a civil war. When he left, he offered her a one-way ticket to Hong Kong and told her that from there he could get her into China. In her words, recorded during a series of interviews I filmed with her in 2003, "I chickened out. Honestly, I thought I'd be a problem for him in a country in turmoil and not speaking the language. And I was scared by the possibility that I would never be able to go back to France."

So her fear of being a burden to him as a foreigner in China and the idea of not seeing her country again were what led her to turn him down.

They continued to communicate even after he returned to China, until early 1950. His letters, which Chris recently uncovered, reveal a generous, wise man who continued to love her and to hope she would accept his offer but who knew it would be too difficult for her. That she kept his letters reveals the depth of her feelings toward him. She told me that she visited the newly opened China in 1982 as part of a US scientific delegation, and even then, when she saw him again for the first time since Dublin, there was still something "special" between them. When the officials told the delegation some cars were waiting to take them somewhere, Peng told them, "No, this one walks with me." She recalled that as the two of them walked, he simply told her, "I'm so glad that you're still wearing sensible shoes."

I took Cécile with me on a business trip to China in 2004, and I had the privilege of accompanying her as she met up once more with Peng, who by then was in his nineties. When we were in his apartment, she opened up her old photo album of Dublin, and I saw his demeanor completely transform: He switched suddenly from communicating formally with her through an interpreter to speaking perfect English, and the two of them disappeared into another world, another time, two old dear friends kidding each other and reminiscing. Later that day, Cécile gave a talk at Peking University, with Peng in attendance. What surprised me the most was that the Chinese lecture attendees all seemed aware of the special relationship between the two of them.

Thank you for writing so nicely about Cécile and for the opportunity to reminisce about this old love story.

Nicolette DeWitt
(nicolette.dewitt52@gmail.com)

Wanting funds to "look everywhere"

David Stevenson's Commentary on the habitable zone as a guide for the search for life in the universe is, as always, trenchant (PHYSICS TODAY, November 2018, page 10). Certainly, we should not focus all of our astrobiology efforts into searches for Earth-like life, lest we miss the variety of life and habitats that may exist elsewhere.

Many proponents of the habitable zone concept never argued otherwise. Rather, they find its value to be not in how it can help us exclude "unhabitable" planets from search efforts but in how it can help us chase the only lead we have in the hunt. Ideally, we would explore all potential habitats for life. But in a funding-constrained environment, it makes sense to allocate resources according to our best guess for where life can be found, with nonzero but smaller efforts spent on unlikely habitats and larger efforts on planets with "naked oceans."

Until "look everywhere" is a funded strategy, spending most of our time in the habitable zone will have to do.

Jason T. Wright
(astrowright@gmail.com)
Pennsylvania State University
University Park

The Heisenbergs and the Goudsmits

Like my father, Werner Heisenberg, I have been the recipient of many questions regarding his role during World War II. After recently rereading Mark Walker's review (PHYSICS TODAY, March 2018, page 55) of David Cassidy's book *Farm Hall and the German Atomic Project of World War II: A Dramatic History*, I would like to reiterate what I personally know about Sam Goudsmit.

He approached me at an American Physical Society meeting around 1978, shortly before his death, and expressed to me how sorry he was about his immediate and strong rebuke of my father at war's end. He felt that our whole family must have been harmed by it. That friendly outreach at the time caught me unaware about the detailed circumstances he referred to. My father's letter on Goudsmit's behalf is a most welcome addition to the factual record. I thank Walker for highlighting it and sharing it with the larger physics community. There was actually such slim hope for Goudsmit's parents, once they were in the horrendous machinery of the Nazi genocide.

My father was an unassuming man with a mind schooled in antiquity (his father was a professor of classics), and he carried the tragedy of the Third Reich within him. He probably accepted that he had tried his best, against great odds, to save Goudsmit's parents. By the same token, he also believed he had done his utmost to prevent Adolf Hitler from having access to a weapon of mass destruction.

The Farm Hall tapes, secretly recorded conversations among 10 captured German scientists including my father, essentially reflect the dense moment of a truth that was irreversible in its consequences for mankind. Great minds are observed as they stumble through that complexity, each from a unique vantage point. No wonder the events at Farm Hall remain a subject of deep inquiry.

Jochen Heisenberg
(j.heisenberg@comcast.net)
Durham, New Hampshire

Correction

May 2019, page 46—In "Microswimmers with no moving parts" by Jeffry Moran and Jonathan Posner, the affiliation of Walter Paxton, Ayusman Sen, Thomas Mallouk, and colleagues should be the Pennsylvania State University. PT

CONTACT PHYSICS TODAY

Letters and commentaries 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 reach us at **http://contact.physicstoday.org**. We reserve the right to edit submissions.

Iron-rich object closely orbits a white dwarf

The newly discovered object could be the core of a planet that survived the transition of its host star into a red giant.

Stars like our Sun can burn brightly for billions of years, generating heat in their cores by fusing hydrogen into helium. When a medium-size star—up to eight solar masses—finally runs out of fuel, nuclear burning no longer drives an outward pressure gradient, and the star starts to collapse due to its own gravity. The contracting star generates a final burst of heat by fusing helium into carbon and oxygen, which causes the outer layers to expand outward.

The star turns into a red giant. In doing so, it engulfs any nearby planets but cannot sustain heat production. Within a billion years, all that remains of the star is the dense (10^9 g/cm^3) stellar core composed of carbon and oxygen. The hot, Earth-sized ember, now called a white dwarf, glows faintly as it cools over tens of billions of years. (See PHYSICS TODAY, March 2019, page 14.)

Theoretical models suggest that asteroids and planets beyond the expanding red giant survive. The largest remaining planetary bodies can scatter the smaller ones into orbits closer to the newly formed white dwarf. Most bodies that wander too close are ripped apart by the dwarf's gravity to form clouds of debris. But the stripped cores of some of those planets, hundreds to thousands of kilometers in diameter, could remain intact if they have sufficiently high internal strength and density.

Observation of objects in orbit around white dwarfs could offer both a preview into the demise of our solar system and a means to directly measure the chemical composition and structure of a planet's inner core. Current technology does not offer a way to directly observe the core of our own planet. But identifying solid bodies that orbit around a dim stellar core, though difficult, is feasible.

Christopher Manser, Boris Gänsicke (both at the University of Warwick), and colleagues have developed a spectroscopic approach with which they have now identified an asteroid-like body orbiting a white dwarf 400 light-years from our solar system.[1] Observations at the largest reflecting telescope in the world, the 10.4 m Gran Telescopio Canarias (figure 1) in La Palma, Spain, made the discovery possible.

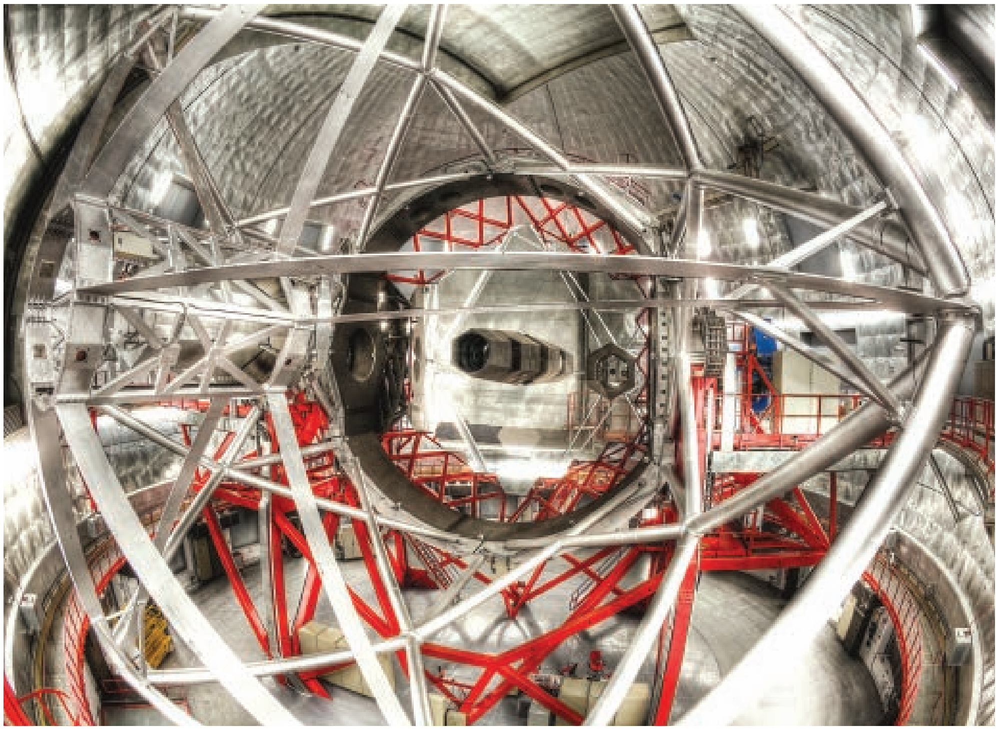


FIGURE 1. THE GRAN TELESCOPIO CANARIAS IN LA PALMA, SPAIN. Shown here is the 10.4 m primary mirror. Spectroscopic observations in rapid succession at the telescope provided evidence of a solid, metal-rich body orbiting closely around white dwarf SDSS J1228+1040. (Image courtesy of GRANTECAN/IAC.)

Gas lighting

The current tally of exoplanets and host stars in the Milky Way exceeds 4000 and 3000, respectively. Most of those discoveries were made with NASA's *Kepler* space telescope using the transit method, which involves identifying periodic dimming as an object passes in front of its host star. (For more on exoplanet detection see the article by Jonathan Lunine, Bruce Macintosh, and Stanton Peale, PHYSICS TODAY, May 2009, page 46.)

Most of those host stars will eventually become white dwarfs.[2] In 2015 Andrew Vanderburg (Harvard University) and colleagues used the transit method to uncover the first direct evidence for a planetary remnant around a white dwarf. Every 4.5 hours, the light from a white dwarf in the constellation Virgo dipped and recovered in a complex pattern as if occulted by several small objects.[3]

Extending the transit method to other white dwarfs has proven difficult. For one thing, it requires that the planetary system's orbital plane lie along the line of sight to Earth. For another, white dwarfs are relatively faint, so the transit technique is limited to those parts of the sky in which other stars are both scarce and dim.

Spectral observations of the residual disks of debris swirling around white dwarfs offer a potentially better tool for detecting planetary remnants. Most disks are composed of dust, as indicated by mid-infrared emissions. In the mid 2000s, Gänsicke had observed emission lines of calcium, iron, and oxygen in the spectra of some disks. He proposed that those lines arise when a white dwarf irradiates the exposed surfaces of rocky planetary debris.[4]

That observation was part of an ongoing project. Manser and Gänsicke had monitored white dwarf SDSS J1228+1040 for 15 years, with the majority of the observations conducted at the European Space Observatory's Very Large Telescope in Chile. From those observations, which made use of the telescope's Ultraviolet and Visual Echelle Spectrograph (UVES) and X-Shooter spectrograph, they created the first detailed image of a white dwarf's debris disk, shown in velocity space in figure 2a. Variations in the observed spectral line shapes provided evidence of ongoing dynamic activity.[5] But it wasn't until the researchers tried in earnest to identify the source of the gas that they made the surprising discovery.

Clockwork calcium

For a more detailed look at the white dwarf and its disk, Manser and Gänsicke turned to the Gran Telescopio Canarias and the OSIRIS imager and spectrograph. By collecting emissions spectra from gases in the debris disk of SDSS J1228+1040 every two to three minutes over several nights in 2017 and 2018, the researchers determined the time it takes material in the disk to orbit the white dwarf.

The spectra contained three bright lines produced by calcium (ii) ions between 850 nm and 866 nm, a signature of metal-rich gases. As expected, each line had two broad peaks, the result of Doppler shifts as the swirling gas in the disk moved toward and away from Earth, oriented at an inclination of 73° as viewed from Earth. But on top of the varying spectrum, light pulsed from one peak to the other every two hours.

Manser explains, "We were originally searching for signs of random variations in the gaseous emission, as we thought gas was being produced by random collisions in the dust." Instead, the periodic variation led the team to conclude that the emissions came from a cloud of gas trailing a likely metallic planetary core that orbited the white dwarf with a two-hour period. Collisions between the planetary core and other surrounding debris could excite the gas as the body sped around its host. The gas cloud trailed the rocky body and boosted one emission peak while the body moved toward Earth and boosted the other while it moved away an hour later.

"Detecting a periodic variation from an emission feature in a white dwarf debris disk is a real first, and characterizing the extra emission as gas from a solid object orbiting close to the white dwarf is very exciting," says Scott Kenyon (Harvard University). Looking for spectral differences in rapid succession offers a new way of detecting solid bodies in any system that shows the calcium ion emission lines.

SDSS J1228+1040 and the seven dwarfs

The satellite's extraordinarily short two-hour orbital period surprised Manser. As shown in figure 2b, the planetesimal orbits 0.73 solar radii (5×10^8 m) away from the white dwarf, which is well inside the debris disk. Adds Jason Nordhaus (Rochester Institute of Technology), "That's pushing up against the boundary of where this object should be torn apart."

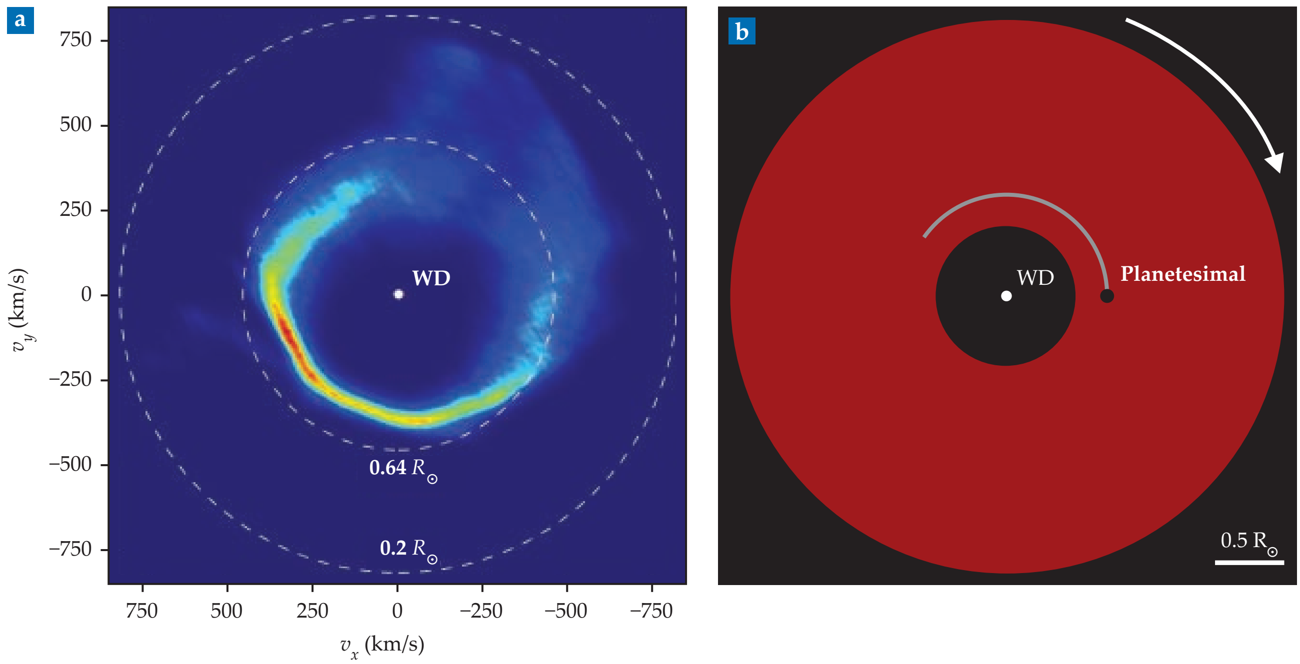


FIGURE 2. THE DEBRIS DISK SURROUNDING THE WHITE DWARF (WD). (a) A map in velocity space shows the pattern of gas swirling in the disk (red is highest flux, dark blue is lowest). The pattern precesses on a time scale of 25 years. The overlaid dashed circles indicate material in orbits at two different distances from the star. The configuration appears inside out because material moves faster in closer-in orbits. The radius of the Sun, $R_\odot$, is 6.96 x 10^8 m. (Adapted from ref. 5.) **(b)** In position space, both the disk and the planetesimal orbit clockwise. The solid red area indicates the region of observed calcium emissions. The gray curved line trailing the planetesimal shows the inferred extent of the gas that generates extra emission. (Adapted from ref. 1.)

A series of calculations led Manser to propose two possibilities for the object's structure. A spherical body as small as tens of kilometers across could be held together by its own gravity provided its density is that of metallic iron, 8 g/cm^3, or higher. Alternatively, an iron-dominated larger body, hundreds of kilometers across, could have a layered internal structure that is strong to avoid being ripped apart. In either case, the original planet would have had distinct layers, like the dwarf planet Ceres. The surviving body could be the iron- and nickel-rich core of a former planet that once orbited much farther away from the star and had its crust and mantle ripped off during the star's explosion. Kenyon observes that "it's interesting to contemplate how the planetesimal got into this mess after having spent most of its previous life far away from its host star."

As of now, astrophysicists know of only seven other white dwarfs that have gas in their disks. Those systems are the next candidates to check for orbiting rocky bodies. Tracking planetesimal behavior over time will help astronomers explain how rocky bodies behave during the final stages of stellar evolution as they form disks around white dwarfs and will also provide the only direct views of planetary inner cores.

Rachel Berkowitz

References

1. C. J. Manser et al., *Science* **364**, 66 (2019).
2. E. Han et al., *Publ. Astron. Soc. Pac.* **126**, 827 (2014).
3. A. Vanderburg et al., *Nature* **526**, 546 (2015).
4. B. T. Gänsicke et al., *Science* **314**, 1908 (2006).
5. C. J. Manser et al., *Mon. Not. R. Astron. Soc.* **455**, 4467 (2016).

Inverted kinetics seen in concerted charge transfer

A counterintuitive phenomenon has now been observed in a new realm.

Just as a round stone rolls faster on a steep slope than on a gentle one, a chemical process speeds up when it's made more energetically favorable. At least, that's what usually happens. But 60 years ago when Rudolph Marcus developed his pioneering theory for electron transfer, he found that in a certain region of parameter space, increasing the driving force—the drop in free energy between the initial and final states—should actually slow the transfer down.[1]

That surprising prediction—the so-called Marcus inverted region—was experimentally confirmed[2] in 1984, and in 1992 Marcus was awarded the Nobel Prize in Chemistry for his theory (see PHYSICS TODAY, January 1993, page 20). Today, Marcus theory is textbook material in chemical kinetics,[3] and inverted regions in electron-transfer reactions are widely observed.

Electron transfer underlies all of oxidation–reduction chemistry, including corrosion, combustion, electrochemistry, and ionic bonding. In photovoltaic cells, the creation and recombination of free electrons and holes are both examples of electron transfer. Engineering a photovoltaic system to reside in a Marcus inverted region can reduce the rate of recombination and enable the extraction of more energy more efficiently.

But electron transfer by itself doesn't involve the making or breaking of chemical bonds that are necessary for manipulating molecular identity or storing energy as chemical fuel. Now James Mayer, Sharon Hammes-Schiffer (both at Yale University), Leif Hammarström (Uppsala University in Sweden), and their colleagues have observed the signature of a Marcus inverted region in a different type of reaction that does rearrange a molecule's atoms.[4] Rather than the transfer of a single electron, their reaction involves the simultaneous transfer of an electron and a proton—that is, a hydrogen nucleus. Such concerted proton–electron transfer is known to be important in biology, solar fuels, and chemical synthesis.

Up is down

Marcus theory stems from the insight that the rate of a charge-transfer process depends critically on what's going on in the solvent or other surrounding medium. In water and other polar solvents, for example, the negatively charged ends of solvent molecules are drawn to positively charged regions of a solute molecule, and vice versa. When charge is redistributed among one or more solute molecules, the energetically preferred solvent configuration changes.

A full representation of the solvent configuration would require a many-dimensional space, but the key features can be collapsed onto a single coordinate, as shown in figure 1. The initial and final states of the reaction have their lowest free energies for different solvent configurations, and in either state, perturbing the solvent away from its preferred configuration increases the free energy in a way that's well approximated by a parabola.

The law of conservation of energy dictates that charge transfer can proceed only when the initial and final states have the same free energy at the same solvent configuration—that is, at the point where their parabolas cross. In most cases, the crossing point is not at the bottom of the initial-state free-energy curve, so it represents a free-energy barrier the system must surmount. The higher the barrier, the slower the reaction.

The seven final-state parabolas in figure 1 represent a series of hypothetical reactions that are identical in every respect except for the free-energy difference between the initial and final states, also called the driving force. (In the lab, such a series might be approximated by chemically altering the electron donor or acceptor to have different affinities for the transferred electron.) In the normal (noninverted) region of parameter space, shown by the blue parabolas, increasing the driving force lowers the crossing point's free energy and speeds up the reaction. But that trend can't continue indefinitely. Eventually, as shown by the gray parabola, the crossing point reaches the initial state's free-energy minimum; in that case, there's no barrier to the reaction, and the charge transfer is as fast as it can possibly be. At still larger driving forces, as shown by the red parabolas, the free energy of the crossing point starts to rise again, and the charge transfer slows down. That's the Marcus inverted region.

Defying diffusion

Early attempts to experimentally observe the inverted region came up short. As the driving force was increased, the rate of electron transfer initially increased, as expected—but then it leveled off and never clearly decreased. The problem was diffusion: In an electron transfer

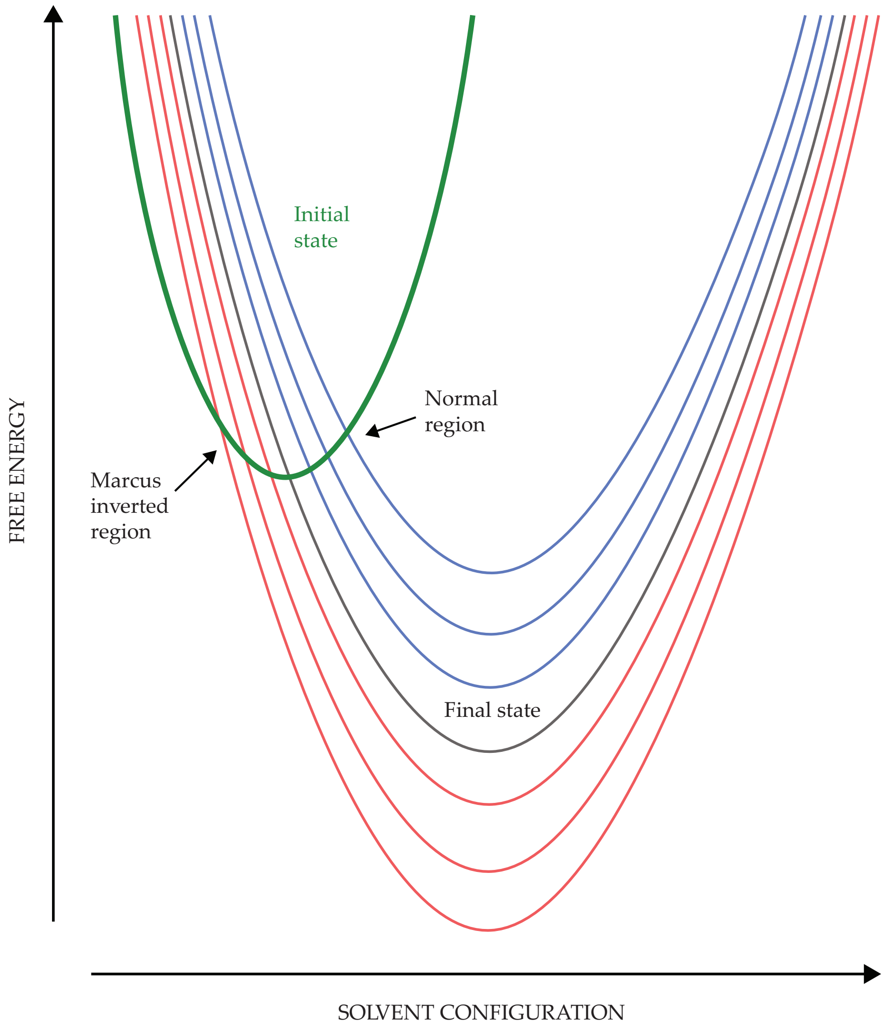
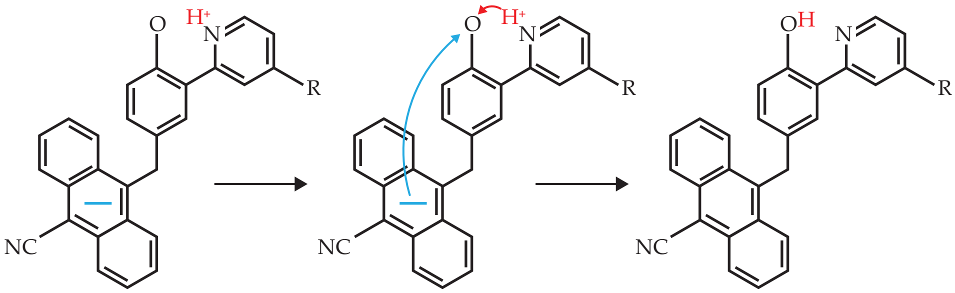


FIGURE 1. ACCORDING TO MARCUS THEORY, a charge-transfer reaction can proceed only at the point where the free-energy curves of the initial and final states intersect. In the normal region of parameter space, lowering the final-state free energy (as represented by the series of blue parabolas) also lowers the free energy of the crossing point, and the reaction speeds up. But in the Marcus inverted region, lowering the final-state free energy (red parabolas) raises the crossing-point free energy, and the reaction slows down.



Charge-separated state | Concerted charge transfer | Ground state

FIGURE 2. IN A THREE-PART MOLECULE called an anthracene-phenol-pyridine, concerted movement of a proton and an electron characterizes the spontaneous relaxation from a charge-separated state to the ground state. Placing different molecular groups at the position marked "R" changes the relative free energies of the initial and final states. The molecules with the larger free-energy changes show slower rates of charge recombination—the signature of a Marcus inverted region.

between two molecules in solution, the measured rate depends not only on the intrinsic transfer rate, as described by Marcus theory, but also on how frequently the donor and acceptor molecules approach each other. When the intrinsic rate is fast, as it is at the onset of the inverted region, the transfer occurs essentially immediately every time a donor and an acceptor get close enough. The rate-limiting step is diffusion, and the intrinsic transfer rate is obscured.

In 1984 John Miller, Lidia Calcaterra, and Gerhard Closs solved the diffusion problem by putting their electron donor and acceptor on the same molecule, connected by a rigid molecular spacer. By guaranteeing that the donor and acceptor would always be in close proximity, they overcame the effect of diffusion and achieved the first unambiguous demonstration of a Marcus inverted region.[2]

In their new paper, Mayer and colleagues also looked at intramolecular charge transfer, this time in a family of three-part molecules called anthracene-phenol-pyridines.[4] Optically exciting an anthracene-phenol-pyridine can give rise to a metastable charge-separated state, as shown in figure 2, with an extra electron on the anthracene unit (the three fused benzene rings in the lower left) and an extra proton, or H^+ ion, on the pyridine (the nitrogen-containing ring in the upper right). To return the molecule to the ground state, the electron and proton both migrate to the phenol unit in the middle. Importantly, the transfer is simultaneous: There's no observable intermediate state in which one of the charges has moved without the other.

The first anthracene-phenol-pyridines were prepared several years ago by Miriam Bowring, then a postdoc in Mayer's group, as part of an effort to push the limits of how fast concerted proton–electron transfer could go.[5] In the unsubstituted molecule (without the CN or R groups in figure 2), she found that the rate of charge recombination was independent of temperature. In a reaction with a free-energy barrier, thermal fluctuations are what push the system over the barrier, so the reaction is faster when the temperature is higher. Temperature independence, on the other hand, meant, tantalizingly, that the researchers had happened upon the zero-barrier reaction that marks the boundary between the normal and inverted regions.

Proton potential

But it wasn't clear that the inverted region would be experimentally accessible. Indeed, theoretician Hammes-Schiffer and her colleagues made the case a decade ago that it shouldn't be.[6] The crux of the argument is that when the H^+ ion moves, it can set a molecular vibration in motion and leave the charge-recombined molecule in a vibrationally excited state. Because a molecular vibration can be approximated by a quantum harmonic oscillator, with a ladder of eigenstates equally spaced over a wide energy range, there's always a state that's close to the right energy for a barrierless reaction. Increasing the driving force, they predicted, should increase the number of vibrational quanta in the final state, with the Marcus inverted region always just out of reach.

Nevertheless, observation of the barrierless reaction was encouraging, and Mayer and colleagues were eager to explore it further. When Bowring presented her results at a conference in Sweden in 2014, their group struck up a collaboration with Hammarström, whose lab was ideally equipped to perform the necessary ultrafast measurements. Giovanny Parada, then a graduate student at Uppsala and now a postdoc with Mayer, also joined the project.

Parada used his synthetic-chemistry expertise to expand the family of anthracene-phenol-pyridines. By attaching different molecular groups at the position marked "R" in figure 2, he could influence the affinity of the H^+ ion for the pyridine and thus tune the charge-recombination driving force over several tenths of an electron volt. Not all the molecules he prepared showed observable charge-separated states, but of those that did, charge recombination was consistently slower in those with higher driving forces. For good measure, the researchers repeated the rate measurements in three solvents of different polarity. They saw the same trend each time.

But what about vibrational excitations—why weren't they blocking access to the Marcus inverted region? It turned out that the charge transfer *was* exciting a molecular vibration, just not with as many quanta as needed to get to the zero-barrier reaction. To see why, Hammes-Schiffer and her student Zachary Goldsmith delved into the quantum details. They found that the wavefunction of the initial charge-separated state had a negligible overlap integral with the vibrational state that would have yielded the zero-barrier reaction. The transition to that state was therefore inhibited. The quantum properties were a consequence of the shape of the potential felt by the proton, so designing molecules with an eye toward that potential could be a route to finding the Marcus inverted region in other concerted charge-transfer systems.

But for now, nobody knows how common or rare the effect might be. The Yale–Uppsala collaboration is on the case, with the theoreticians exploring large regions of parameter space to guide the experimenters' next choice of molecules to study. "We'd like to think that the phenomenon will prove to be widespread," says Mayer, "because then it could be used in more complex systems to tackle challenges such as solar-energy conversion." Because concerted proton–electron transfer is common in biology, in processes such as photosynthesis and respiration, another intriguing question is whether nature already exploits the Marcus inverted region in biological pathways. If so, understanding the inverted kinetics could be key to mimicking those functions in synthetic systems.

Johanna Miller

References

1. R. A. Marcus, *Discuss. Faraday Soc.* **29**, 21 (1960).
2. J. R. Miller, L. T. Calcaterra, G. L. Closs, *J. Am. Chem. Soc.* **106**, 3047 (1984).
3. See, for example, P. L. Houston, *Chemical Kinetics and Reaction Dynamics*, Dover (2006), p. 155.
4. G. A. Parada et al., *Science* **364**, 471 (2019).
5. M. A. Bowring et al., *J. Am. Chem. Soc.* **140**, 7449 (2018).
6. S. J. Edwards, A. V. Soudackov, S. Hammes-Schiffer, *J. Phys. Chem. B* **113**, 14545 (2009).

Flows of volcanic rock and gas ride a carpet of air

The two-phase fluid drives hot gas to its base to lower frictional forces.

FIGURE 1. THE VOLCANO ERUPTION SIMULATOR FACILITY at Massey University in New Zealand produces experimental pyroclastic flows—hot avalanches composed of air, ash, and rock—that allow the study of their dynamical behavior. In this series of images spanning about a second, the bottom meter of a two-layer flow passes a fixed observation point. The arrows denote the boundary between a denser particle layer overridden by dilute, turbulent ash.

Mount Vesuvius erupted in 79 CE with little warning. The volcano's pyroclastic flows—hot avalanches composed of air, ash, and rock—obliterated everything in their path. The Roman cities of Pompeii and Herculaneum and the remains of about 1500 people were later excavated from under several meters of ash. In contrast to the viscous, ambling lava flows of other volcanoes, such as Mauna Loa in Hawaii, pyroclastic flows cruise across land at speeds of about 10–30 meters per second for tens of kilometers without slowing down.

Pyroclastic flows have long perplexed volcanologists. Given volcanic particles' high static friction, they should stay put on slopes of up to 35–45°. But scientists have observed pyroclastic flows traveling over land surfaces with average slopes of just 8° and sometimes even upslope for short distances.

Volcanologists can't easily measure the physical properties of pyroclastic flows in the environment because of the danger to people and the destruction of field instruments. Instead, researchers turn to numerical models. For a few decades, high pore pressure—that is, the pressure in the space between particles—was suspected of modifying frictional forces. But without a clear understanding of how pyroclastic flows generate and sustain pore pressure, scientists couldn't conclusively test the hypothesis. Rather

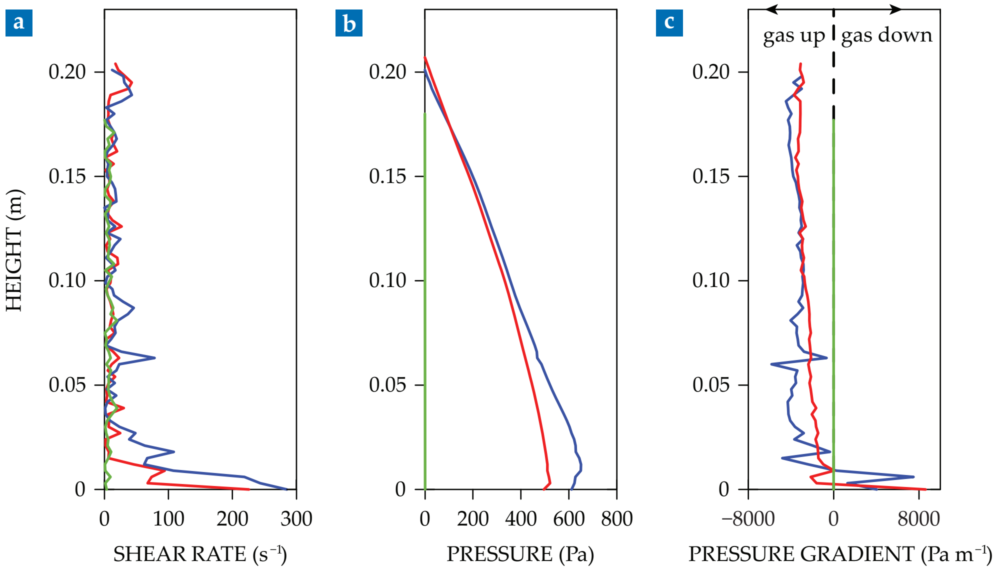


FIGURE 2. PYROCLASTIC FLOWS glide across the ground surface on hot air. The blue, red, and green lines denote the flow properties at 380 ms, 690 ms, and 1290 ms, respectively, at a static observation point. The measured shear rate **(a)** varies with the height of the flow by three orders of magnitude and **(b)** forms a pressure maximum a few centimeters above the flow's base. The resulting pressure gradient **(c)** moves air downward toward the lower-pressure area, effectively lowering the particle concentration and decreasing friction so that pyroclastic flows can travel at high speeds for long distances.

than explicitly simulate the dynamical evolution of pyroclastic flows, modelers had to resort to using empirically derived friction coefficients to fit models to observations.

Now Gert Lube of Massey University in New Zealand, Eric Breard of the University of Oregon, and their colleagues have identified the source of pyroclastic flows' friction-defying capability. In New Zealand, Lube oversees a laboratory facility known as PELE—for pyroclastic flow eruption large-scale experiment—that can safely create large flows. The researchers discovered a low-pressure, dilute air layer at the base of pyroclastic flows that reduces internal and ground-surface friction.[1]

Paying tribute to PELE

Before PELE, researchers built miniature volcanic eruptions in the lab. Volcanologists at the University of Clermont Auvergne in France sent volcanic ash deposits down a 3 m flume to measure particle velocities.[2] At the University of Bari Aldo Moro in Italy, gas and ash erupted out of a human-made volcanic chimney.[3] Those and other, similar experiments have helped uncover the turbulent dynamics of pyroclastic flows, but challenges in scaling from real life to the benchtop have limited what scientists can learn about the friction-defying capability of pyroclastic flows. To understand how flows modify friction on a downscaled system, says Lube, "you need to make sure the balance of forces is on the same range as they would be in nature."

Another problem that researchers have had with laboratory-scale eruption analogues is with measuring certain flow properties, because they use less material over shorter distances than nature. The pressure in the flow, for example, can diffuse too fast for sensors to measure it. After some earlier small-scale efforts, says Lube, "it was clear in the volcano community that we needed to do our experiments at a large scale."

So in 2011 Lube and some collaborators began designing a larger facility at Massey. Breard joined the team as a graduate student in 2012 when the construction started. PELE operates by combining air, heated up to 130 °C, with 1000–1300 kg of rock and ash collected from the last eruption of New Zealand's Taupo volcano, more than 1800 years ago. The mixture falls freely down a 9 m vertical column to a chute that creates 10-m-high pyroclastic flows that can run horizontally for 25 m. Using high-speed cameras, pressure sensors, and load cells along the chute, Lube and Breard can capture the time-evolving velocity, pore pressure, and mass of the experimental flows.

"We were surprised when these flows moved out like a liquid," says Lube. "We knew from that moment that we have this process, and now we just need to look inside and measure it." By assuming that a flow's mass and volume are conserved over distances of a few centimeters, Lube, Breard, and their colleagues could analyze variations in the flow's kinetic and potential energy by measuring its height and velocity. With that information, they calculated the time-varying effective friction coefficient that represents both the internal friction between particles in the flow and the friction between the flow and the bottom of the chute.

PELE produces two-layer pyroclastic flows consisting of a denser particle layer overridden by a dilute, turbulent layer. The series of images in figure 1 show the bottom meter of an experimental flow that passes a stationary point; the arrows point to the boundary between the denser layer with a high particle concentration and the overriding turbulent ash cloud. Lube, Breard, and their colleagues found that the friction acting on and internal to the particle layer varied with depth. The friction coefficient was 0.2–0.3, in agreement with natural volcanic deposits. The researchers observed that a third, basal layer developed in the lower third of the particle layer and had even smaller values of 0.05–0.21. The smaller values are possible because of the low particle concentration in the basal layer.

A mechanism revealed

Breard recounts an idea from coauthor Jim Jones, a chemical engineer at Massey, that led the team to consider how an air layer could act as a lubricant for a pyroclastic flow: "When we discussed changes in particle concentration, he thought we needed to discuss how shear changed through the flow." A pyroclastic flow tends to behave like a fluid. Normally in a fluid, the air moves parallel to the flow direction, and high pressure at the base eventually forces the less dense air up and out of the two-phase flow, stopping the fluid in its tracks. But in a pyroclastic flow, shearing can stall that process.

If the shear rate changes with depth and keeps the air longer in the flow, the pressure gradient can change too. In that situation, a low-pressure air layer develops at the base and drives air downward while solid material hovers above. As figure 2a shows, the shear rate varies by three orders of magnitude across the depth of the experimental pyroclastic flows and is highest at the base. Those shear rates generated a pressure maximum a few centimeters above the base, as can be seen in figure 2b.

In particle–gas flows, the gas responds more quickly to pressure gradients than the particles do. As the pressure gradient drives the air toward the lower-pressure basal layer as shown in figure 2c, it effectively decreases that layer's particle concentration and, consequently, the friction between the flow and the ground surface. The air moves faster than the particles, so the pyroclastic flows can ride the air for long distances without slowing down. The mechanism also is self-reinforcing: Less friction lowers the deceleration of the flows. Lower deceleration perpetuates the higher shear gradient at the base, which drives the air downward.

"The role of air was proposed decades ago based on studies where air is injected into the base," says Olivier Roche, a volcanologist at the University of Clermont Auvergne. The novelty of this study, he says, "is the demonstration that pore-gas pressure can arise in the experimental flow." Such results would not have been seen without the high velocities and shear rates produced by the large-scale experiments.

Now that volcanologists understand the main physical process for how pyroclastic flows affect friction, natural-hazard modelers may be able to develop better simulations. "More complex models showing the formation of pore pressure and how long it can be sustained should be more accurate" than earlier models, says Breard. Where a pyroclastic flow goes depends mostly on the topography, but Breard says that new models that incorporate the air-layer mechanism may be able to make more precise estimates of the distances that pyroclastic flows travel.

Volcanoes in Chile, Nicaragua, and other places in Central and South America pose dangerous hazards. The 3 June 2018 eruption of Volcán de Fuego in Guatemala produced pyroclastic flows that buried nearby towns, killing and injuring dozens of people. With improved numerical models, "you can train people to use them and convey to people the information in the danger area," says Breard. "Those hazard models are used by researchers who ought to ensure that populations living in these volcanically active areas know about them and their predictions."

Alex Lopatka

References

1. G. Lube et al., *Nat. Geosci.* **12**, 381 (2019).
2. L. Girolami et al., *Bull. Volcanol.* **72**, 747 (2010).
3. P. Dellino et al., *J. Geophys. Res. Solid Earth* **112**, B04202 (2007).

A metamaterial solves an integral equation

By iteratively processing an optical signal, the structure functions as an analog computer.

As computers' processing speeds grow ever faster, they are approaching the limits of their electronic components. More speed requires additional power and smaller, denser chips, but miniaturization inhibits heat dissipation. Fiber-optic cables already transmit information encoded in light faster and more reliably than electrical cables can carry electronic signals. If computers also processed information as light instead of electronically, they could boost computational speed and mitigate the excess heat generation that besets electronic circuits as they become faster and smaller.

Switching from electrical to optical computing has other clear advantages.[1] In addition to faster signal propagation and less heat generation, optical signal processing is also inherently parallel. To understand why, consider imaging with a lens. The light intensity in the back focal plane is the spatial Fourier transform of the original signal. Performing the same Fourier transform electronically would require the function value at each point to be calculated sequentially, but with optics all points are evaluated simultaneously.

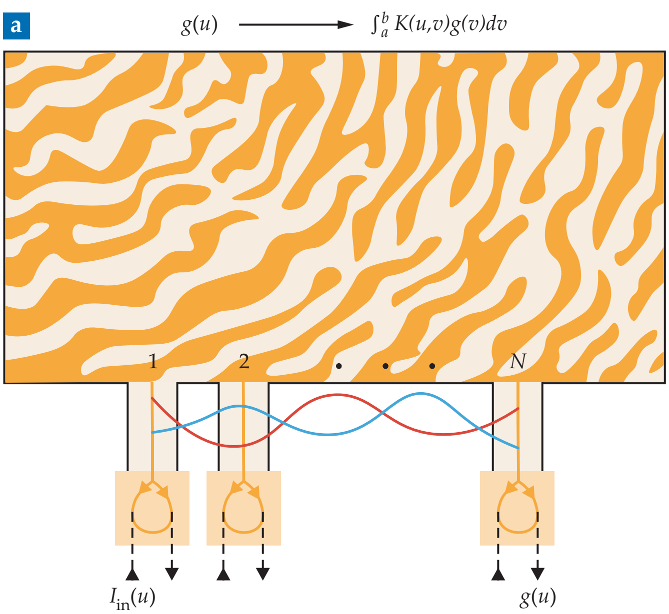
Researchers are developing materials and devices that may one day lead to practical optical computing technologies. Many of them manipulate light in the same way that switches and logic gates direct electronic signals. Those switches and gates would then be assembled into larger, multifunctional networks. But there's no reason the strategies used in light-based computers have to mimic those of conventional computers, and other approaches target devices designed to perform specialized functions in ways that could surpass electronics.

Nader Engheta at the University of Pennsylvania and his collaborators have now built a device that solves integral equations using light.[2] Their metamaterial block repeatedly manipulates a microwave input signal until it reaches a steady state that represents the equation's solution. Although the device is not yet reconfigurable or programmable, it is smaller than those using other optical processing schemes and has the potential to solve integral equations much faster than a conventional computer.

Designing a device

Engheta first proposed using metamaterials for electromagnetic signal processing[3] in 2014. He envisioned layered blocks whose varying permittivity and permeability would perform a mathematical operation on the incoming signal such as taking a derivative or integrating. Other groups have used different strategies to tackle optical signal processing,[4] but metastructures have the potential to form smaller devices and be incorporated into circuits. They also process signals nonlocally: The input signal scatters through the device and affects the output at every point in the function's domain, which allows the metastructure to solve global problems like integrals.

The new device developed by Engheta's group does more than perform a



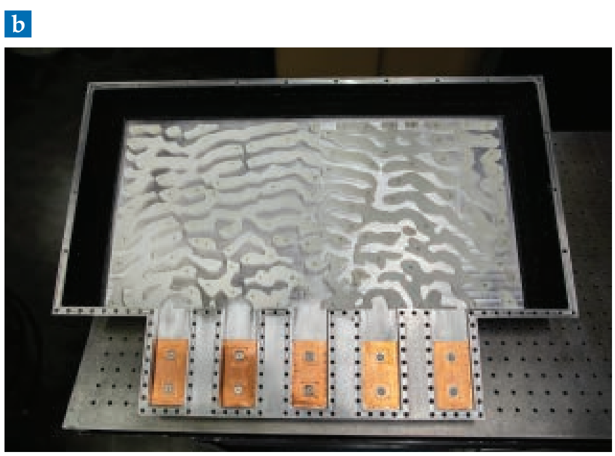


FIGURE 1. A METASTRUCTURE DEVICE INTEGRATES an incoming signal over a kernel K to solve an integral equation. **(a)** The integral over K is encoded in the metastructure's pattern, illustrated here by the wavy orange pattern. Waveguide coupling elements (orange squares) introduce an input signal $I_{in}(u)$, for $u = 1, \ldots, N$ (red curve), that moves through the device and then returns to the coupling elements. A fraction of that signal is removed to track the device's progress. The rest is directed back into the metastructure, and the process repeats until the output signal (blue curve) reaches the steady-state solution $g(u)$. (Adapted from ref. 2.) **(b)** The microwave-scale device is about 30 cm by 60 cm and uses a metamaterial of polystyrene and air. (Photo courtesy of Eric Sucar.)

mathematical operation—it solves equations. The researchers focused their attention on equations of the form

$$g(u) = I_{\text{in}}(u) + \int_a^b K(u,v)g(v)dv.$$

Known as Fredholm integral equations of the second kind, they are linear integral equations that arise in many fields of science and engineering, such as antenna theory and quantum perturbation theory.[5] The function $g(u)$—the quantity to be solved for—appears on both sides of the equation, so the researchers applied an iterative strategy. The input signal $I_{\text{in}}(u)$ passes through a metastructure that integrates the signal over a kernel, $K(u,v)$, whose form reflects the integral's underlying physics. The integrated signal leaving the metastructure is directed back in as the input for the next integral step. Eventually the output signal reaches a steady state, which is the solution g. The approach works in principle because the wave scattering in the metamaterial is mathematically equivalent to multiplication between a matrix operator and an input vector.

Figure 1a shows a schematic of the device. The large, patterned region represents the metastructure and its varying dielectric constant. The N waveguides at the bottom direct electromagnetic signals into and out of the metastructure. Each waveguide contains a coupling element (orange squares) with two coaxial cables, one to introduce I_{in} and another to track the device's progress by measuring the output signal g. The researchers monitor the process using only a small fraction of the signal that returned to the coupling elements; the rest of the signal is directed back into the metastructure for repeated processing. Each waveguide handles a discrete value of u in the integral's domain.

The instructions for evaluating an integral over a particular K are encoded in the metamaterial's structure. In particular, the researchers used an algorithm to translate a scattering matrix corresponding to the integral into a dielectric-constant distribution for the metastructure. But the fact that the kernel is hardwired into the device doesn't mean the device only solves one equation—changing I_{in} produces a different solution g.

After Engheta and colleagues confirmed through simulations that their strategy works, they built the device shown in figure 1b that solved equations using microwaves of wavelength $\lambda = 6.85$ cm in air. Although the simulated metastructures had continuously varying dielectric constants, for simplicity the real device contained only two materials—polystyrene and air—and used kernel structures that Engheta described as "Swiss cheese." But there's nothing special about those materials. Selecting different or multiple dielectrics would just change the geometry of the device.

The researchers tested their device with an input signal at each of the five waveguides and measured the amplitude of the steady-state output. Their results, shown for the center waveguide in figure 2, matched expected theoretical results and simulations. Although using only five waveguides simplified the experiments, the coarse discretization caused some deviation from the theoretical result. That problem is not inherent to the technique though: The same metastructure design also works with more waveguides, and simulation results using $N = 20$ match the theory more closely.

For a simulation of an input signal propagating through the device, see the online version of this story.

Adaptations and reconfigurations

Optical computing has a long way to go before it challenges electronic computing. "One big advantage of current computers is that they're programmable," says Engheta. "The piece of computer that you have in front of you is one hardware but you can do many things with it." To be programmable, the structures in an optical computer will have to be reconfigurable. The group is exploring phase-change disk technology used for rewritable CDs to make reconfigurable metastructures that can then serve as more than one kernel.

Computing with light could also relax the requirement that functions be discretized for numerical analysis. In their proof-of-concept device, the researchers did discretize their function by using a finite number of waveguides because it was a convenient way to control I_{in} and measure g. "But the concept is not necessarily limited by discretization," says Engheta. In principle, a device illuminated by a continuous wave could perform the same function. The metamaterial would work in the same way,

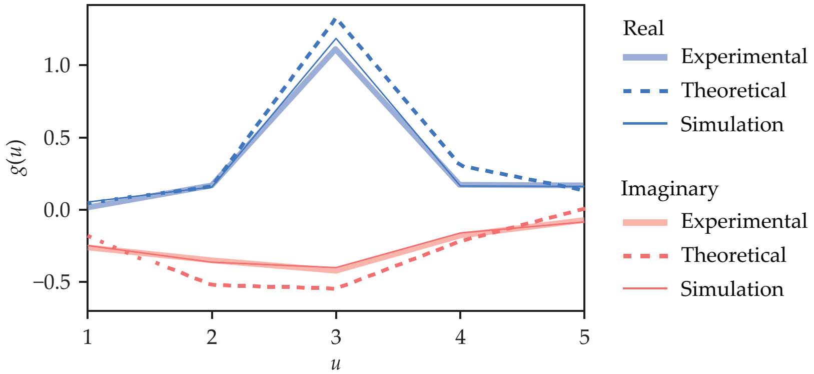


FIGURE 2. AN INPUT SIGNAL I_{in} = [0 0 1 0 0] that was introduced through the center waveguide of a five-input device produced a steady-state output signal $g(u)$ at all five waveguides. The experimental measurements (bold lines) of the real (blue) and imaginary (red) parts of the electric field closely matched theory (dotted lines) and simulations (thin lines). (Adapted from ref. 2.)

but the researchers would need to devise a new strategy for introducing I_{in} and measuring g.

The device that the researchers built was the size of a briefcase—about 30 cm by 60 cm—also for convenience, which meant that they had to use microwaves. It's easier to engineer a device on that scale than on the microscale. And polystyrene, which is commonly used with microwaves, is readily available and inexpensive. Now that the researchers have demonstrated a proof of concept, they want to shrink the device so that it works in the near-IR. Features and even whole devices would then be on the micron scale and potentially suitable for chip-based technology.

Moving to the near-IR will also improve the device's speed. The group's analysis shows that it takes about 300 times the wave period for the device to converge on a solution. The microwave frequency in the group's experiments is a few gigahertz, so it takes tens of nanoseconds to solve an equation. In the near-IR that time would be picoseconds—faster than current processors execute a single instruction.

Now that they've constructed a device that can solve one type of integral equation, Engheta's group is working to widen its applicability. In addition to solving different forms of linear integral equations, they would also like to introduce nonlinearity and combine multiple swiss-cheese kernels to solve systems of coupled equations.

Christine Middleton

References

1. H. J. Caulfield, S. Dolev, *Nat. Photonics* **4**, 261 (2010).
2. N. M. Estakhri, B. Edwards, N. Engheta, *Science* **363**, 1333 (2019).
3. A. Silva et al., *Science* **343**, 160 (2014).
4. X. Lin et al., *Science* **361**, 1004 (2018); T. Zhu et al., *Nat. Commun.* **8**, 15391 (2017).
5. A.-M. Wazwaz, *Linear and Nonlinear Integral Equations: Methods and Applications*, Springer (2011).

PT

Voltage Controlled Current Source

Your science. Our tools. SRS.

- Sources/sinks AC & DC current
- Analog voltage control input
- 1 nA/V to 50 mA/V gain
- 200 kHz bandwidth
- ±50 V compliance
- Ultra-low noise design
- RS-232 & optical fiber interfaces

The Model CS580 Voltage Controlled Current Source creates new capabilities for researchers needing an ultra-low noise current source in a flexible, easy to use instrument.

The CS580 is a natural companion to sensitive AC instruments such as lock-in amplifiers, providing a easy way of generating precision current directly from an AC or DC control voltage.

The CS580 is a welcome addition to any research lab studying semiconductors and transport phenomena, superconductivity, or nanotechnology, to name just a few.

Model CS580 ... $2795 (U.S. list)

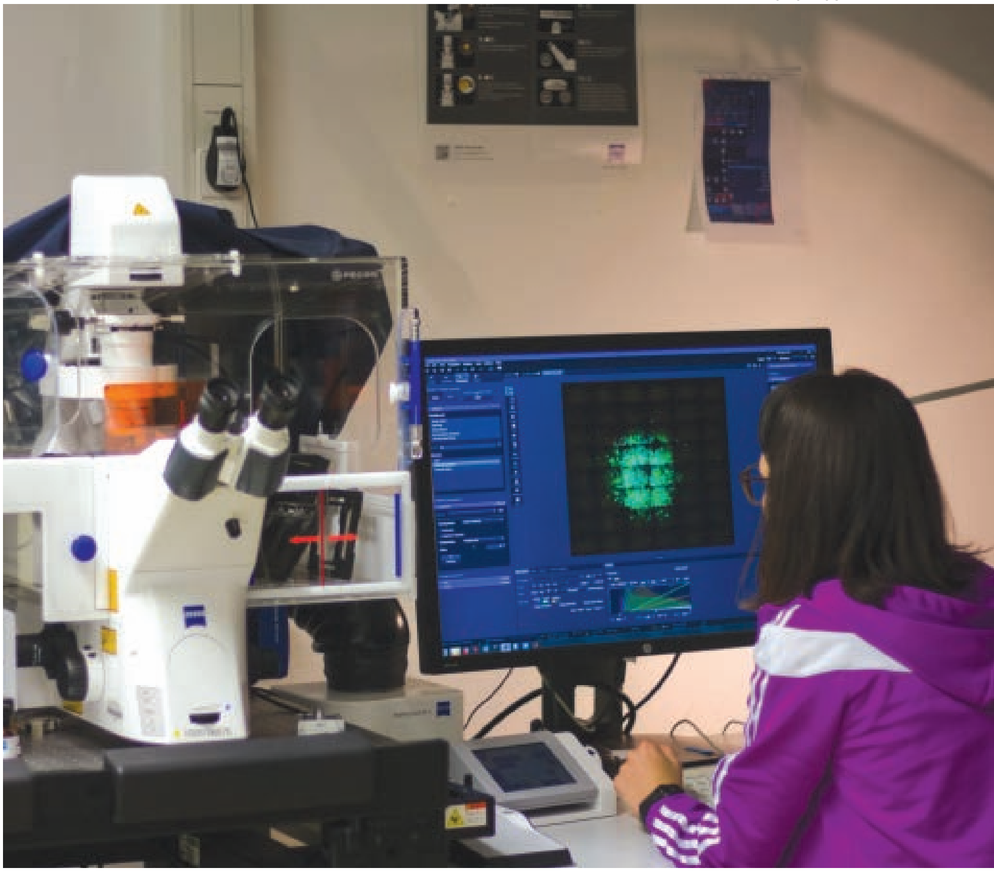
Stanford Research Systems
www.thinkSRS.com • Tel: 408-744-9040

www.thinksrs.com/products/cs580.html

Physical and mathematical approaches yield insights into how cancer develops and spreads

Whether cancer cells survive, grow, or remain dormant depends in part on mechanical forces and other physical phenomena.

HANS KUBITSCHKE



ELIANE BLAUTH, a master's student in Josef Käs's lab at the University of Leipzig in Germany, looks at an image of a fluorescently labeled tumor.

Despite improved prognoses for certain types of cancer, progress in combating the disease has been modest since the declaration in 1971 of the "war on cancer." Heart disease and cancer are the leading causes of death in the US. And with the nation's population aging, the number of cancer deaths is expected to continue rising. Aside from cancer's lethality and enduring mystery, the realization that cells respond to physical and mechanical cues in addition to chemical and genetic ones is drawing physical scientists to study the disease.

Historically, physicists have played a major role in developing diagnostic and treatment tools for cancer—from x rays and MRI to lasers, radiation therapies, and designer drugs. But in the last decade or so, physical scientists and engineers have increasingly delved into understanding the disease, an area that had been almost exclusively the domain of biologists and medical researchers.

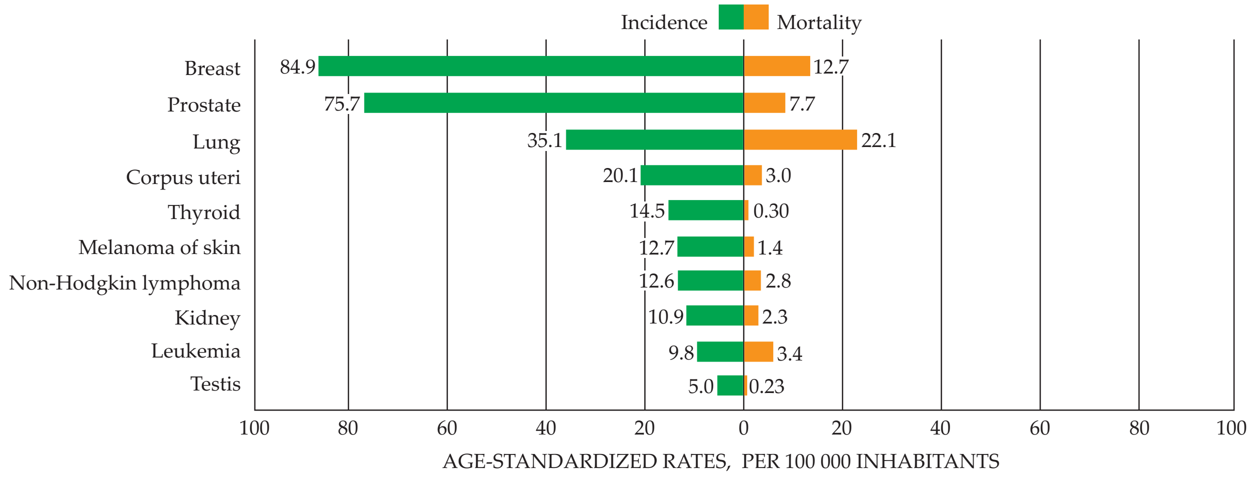
By helping to select the most relevant data, bringing more math to analyze them, and creating meaningful theoretical constructs, physical scientists have already contributed significantly to understanding cancer, says Anna Barker, who a decade ago at the National Cancer Institute (NCI) was key in launching the Physical Sciences–Oncology Centers (see PHYSICS TODAY, November 2014, page 22) and is now codirector of the Complex Adaptive Systems Network at Arizona State University. Those advances are on the brink of being powerful, she says. Physical sciences and information theory are "the new frontier of cancer research."

Physical scientists study myriad aspects of cancer, including how tissue stiffness correlates with cancer; how cancer cells modify their environment, differ from healthy cells, and migrate, select, and create sites for metastasis; whether proteins act as mechanical–chemical switches; and how to improve prognosis predictions. Researchers use mathematical and computational models, machine learning, engineered systems, tumor cells in two- and three-dimensional cultures, animal subjects, and clinical studies with humans. The wide-ranging studies help develop a fuller picture, says Josef Käs, a biophysicist at the University of Leipzig in Germany. "Cancer is such a complex problem that there is no single $100 million question. But we are understanding more by the day."

Cancer repurposes healthy cells

One point of entry for physical scientists who want to understand cancer is tumor progression. Kandice Tanner, a physicist at the NCI, uses zebrafish to study organ selectivity in cancer metastasis (see the image on page 28). Her aim is to identify the physical properties that drive the selection of cancer sites and how the microenvironment influences whether cells survive, grow, or remain dormant. Sometimes having been trained outside of biology brings a fresh perspective, she says, "but you have to be careful that ideas are really applicable." (See also the interview with Tanner about her research and career trajectory at http://physicstoday.org/tanner.)

Xavier Trepat, a group leader at the Institute for Bioengineering of Catalonia in Barcelona, Spain, focuses on the forces involved in cancer development. In



INCIDENCE AND MORTALITY RATES in the US for different cancer types. (Reproduced with permission from J. Ferlay et al., *Global Cancer Observatory: Cancer Today*, International Agency for Research on Cancer, 2018. Available from the interactive website https://gco.iarc.fr/today, accessed 14 May 2019.)

metastasis, for example, cells detach from a primary tumor, move through tissue, travel through the bloodstream or lymphatic system, and eventually form secondary tumors. Each step involves many proteins and genes, says Trepat, but the steps also involve mechanical processes. To enter a blood vessel, a cancer cell has to push other cells out of the way and deform itself. His lab does *in vitro* experiments to measure the forces in such processes.

The 1–100 nN forces exerted by healthy cells enable motion, and they also transmit biological signals—telling a cell to secrete a protein, express a gene, divide, die, or perform another action. Cancer cells commandeer those signals to compel other cells to do what they need.

In his experiments, Trepat mixes tumor and healthy cells, "because any function you look at in a tumor is affected by nontumor cells." He and his team found that skin-cancer cells are not themselves mobile. Rather, they exploit the mobility of healthy fibroblasts, cells that secrete collagen and other macromolecules to form connective tissue and facilitate wound healing. Cancer cells adhere to fibroblasts and get pulled from the tumor. "These tumor cells hijack the function of fibroblasts to help them migrate," he says. "This cross talk between healthy and cancerous cells ultimately helps cancer spread."

Many tumor types are stiffer than normal tissue; cancer is often diagnosed by palpation. Tissue stiffens when the fibers in it become more cross-linked or stretched, or it becomes denser through proliferation of cancer cells. As a tumor grows, it increases the local pressure, which cancer cells may better withstand than normal cells can.

Tissue stiffness can be measured *in situ* using ultrasound or magnetic resonance elastography. And Viola Vogel and her group at ETH Zürich recently synthesized a peptide that binds only to relaxed tissue fibers. That selectivity allows the researchers to distinguish—for the first time, Vogel says—relaxed from stretched fibers. "Surprisingly, we found that tumor tissues contain a large fraction of relaxed fibers. Since nothing is known about the tension of tissue fibers in healthy and diseased tissues, we need to figure out what that means physiologically." She adds that the peptide might be used to identify and target diseased fibers remaining after chemotherapy.

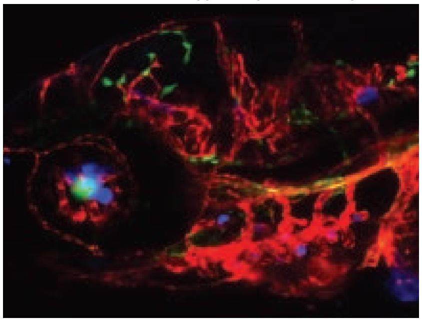
Tumors and nearby healthy tissue can be excised and then compared when stretched and strained. Liver-cancer cells, for example, proliferate and lose their liver characteristics more when they are grown on stiff substrates than on softer ones. Dennis Discher, a biophysicist at the University of Pennsylvania who heads one of the NCI's 12 Physical Sciences–Oncology Centers, says that by the 2000s, researchers understood that a tissue's softness is important for its health. The risk of liver cancer rises if the organ stiffens, whatever the cause—alcohol, a virus, cirrhosis, fibrosis, or other source.

Stiffness in the extracellular matrix drives cancer progression, says Janine Erler, a cancer biologist at the University of Copenhagen in Denmark. She and physicist Lene Oddershede of the university's Niels Bohr Institute embed healthy and cancerous cells in collagen matrices and use optical tweezers to measure the cell stiffness as a function of tissue density. The matrix density doesn't alter healthy cells, says Oddershede, but invasive cancerous cells can become softer or stiffer in response to the stiffness of the matrix. "In a stiffer microenvironment, cells behave more aggressively," Erler says. What's not known is whether the matrix stiffness is a driver or a by-product of the aggressiveness.

Only aggressive cancerous cells can squeeze through pores in the surrounding tissue matrix. The pair's goal is to understand what regulates cell behavior—what cues cells to invade or renders them unresponsive to drugs, says Erler. "If we can understand how a cell senses it's in a stiff environment, we may be able to change the response. It opens the possibility that we could trick cells into thinking they are in a soft environment."

Käs and his group in Leipzig observe tumor chunks that have been surgically removed from people with breast and cervical cancers. They have seen areas where more than half the cells are round and don't move. In between are fluid areas where cells stream to the surface and leave the tumor. It's like the jamming–unjamming transition in colloids, he says. "Cells can jam and unjam through shape changes. What is becoming clear is that it's not what the single cell does, it's collective behavior." (See, for example, the PHYSICS TODAY articles by Robert Evans, Daan Frenkel, and

COLIN PAUL AND KANDICE TANNER



HUMAN BREAST CANCER CELLS were injected into a zebrafish, where they spread and colonized in the brain (blue).

Marjolein Dijkstra, February 2019, page 38, and by Jasna Brujic, November 2010, page 64.) In pathology studies, Käs adds, the presence of a lot of round cells in a tumor correlates with a better prognosis; the presence of elongated cells suggests metastasis.

Killer metastasis

Generally, a patient can be cured only when the primary tumor can be fully removed before metastasis occurs. Even then, cancer cells may have spread undetected and the disease may recur. For Robert Austin, a Princeton University physicist, the most pressing question is how cancer metastasizes. "Ninety percent of cancer deaths are due to metastatic cancer," he says. "That is the linchpin." (See the article by Chwee Teck Lim and Dave Hoon, PHYSICS TODAY, February 2014, page 26.)

Tumors are highly stressed environments: acidic, hypoxic, and low in nutrients. Austin and his group image cell movement on microfabricated silicon substrates while varying conditions such as drug gradients and fluid flow. They mix metastatic cell lines "to get as generic as possible," he says.

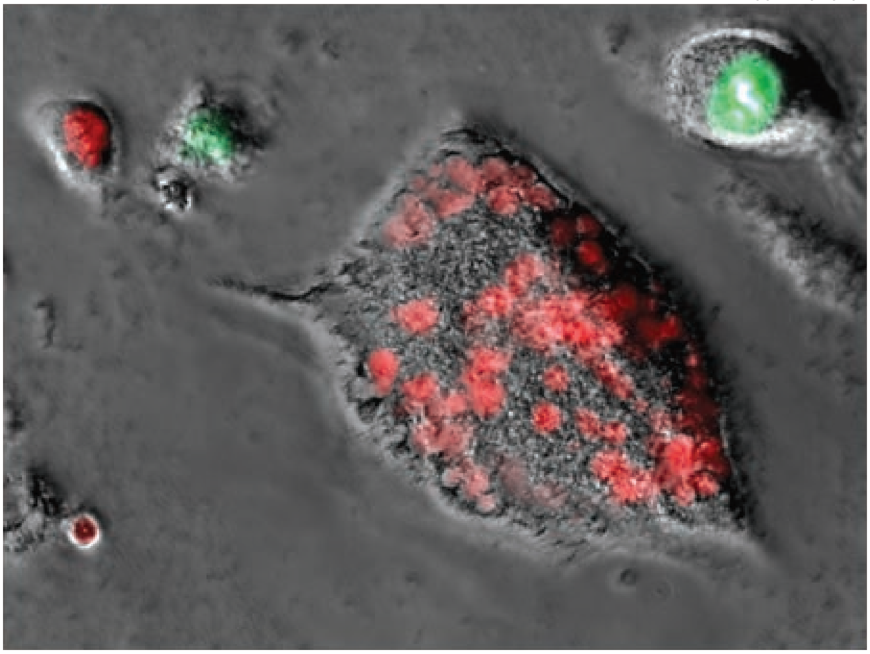
The cells that survive the harshest conditions become polyploid—they engulf other cells and end up with multiple sets of chromosomes from different cell types (see the image on page 29). "It's no longer a simple prostate or bone marrow or breast cancer cell," says Austin. The entanglement is a survival strategy that "gives cancer cells a bigger phase space to search for how to solve a problem." A new avenue of therapy, he suggests, could be drugs that target cells with polyploid nuclei and the processes through which the nuclei become polyploid.

Cynthia Reinhart-King of Vanderbilt University's department of biomedical engineering also uses fabricated systems to study cancer. "We can engineer collagen to look like tumor collagen, and we can adjust the pore size and the fiber size of collagen." With such systems, she and her group try to find out, for example, how cells move along fibers and why tissue density is a prognostic indicator for breast cancer. "We tailor systems to look at aggressive cells, the role of tissue structure, and what drives cells to move and grow. We study single cells and collective behavior."

One of the biggest contributions by physicists and engineers, Reinhart-King says, is showing that it's not just the cells themselves that promote disease. "We have identified that tissue structure, stiffness, porosity, and fluid gradients can all contribute to cancer progression." In addition, she says, researchers have discovered that conditions in the microenvironment can make normal cells behave like cancer cells and coax cancer cells to behave like healthy ones.

Metastasis involves leader–follower

GONZALO TORGA



THIS METASTATIC PROSTATE CANCER CELL has more than 30 sets of chromosomes, accumulated from a combination of engulfing foreign nuclei and exposure to a chemotherapy drug. The red dots are fluorescently labeled proteins in the nuclei.

behavior, with some cells showing the way in leaving the primary tumor, says Reinhart-King. She and other researchers have found that energy determines which cells lead, and when those cells run low on energy, others take over. But how individual cells become leaders and followers in the first place remains unclear. And given that a primary tumor can shed thousands of cells a day, why are there relatively few metastatic sites?

Predicting prognosis

Herbert Levine, a theorist at Northeastern University, analyzes signaling and decision making in cancer cells. (See the article by Jané Kondev, PHYSICS TODAY, February 2014, page 31.) What drives cells to send out particular signals? How do they decide to move? The cytoskeletal protein vimentin, for example, is present only in cells that can move, and a measurable change in the protein's concentration can indicate when a cancer cell becomes motile, and potentially metastatic, Levine explains. He might start his dynamical models with a handful of variables and then increase the number to incorporate suggestions from cancer biologists. "We try to understand how a cell's decision making correlates with different parameters," he says. However, the results are "squishy. They are not exact answers. But the good thing is that they may be generalized to various cell lines."

Theoretical physicist Benjamin Greenbaum, director of the Center for Computational Immunology at Mount Sinai, studies immune-driven tumor evolution and tumor response to therapies that activate or suppress the body's immune system to fight cancer. Why, for example, does immunotherapy do nothing for one group of patients with metastatic melanoma, but in another it seemingly wipes the disease out? He is also part of an interdisciplinary team, funded by the charity Stand Up To Cancer, that is investigating why a small fraction of patients survive pancreatic cancer for an unusually long time. He and colleagues model the evolutionary dynamics of tumors. They compare a tumor's molecular alterations, such as arise from mutations, that the immune system can recognize, with the aim of pointing the way to treatments for nonresponders. (Last year's Nobel Prize in Physiology and Medicine was awarded to James P. Allison and Tasuku Honjo "for their discovery of cancer therapy by inhibition of negative immune regulation.")

Clare Yu, a theoretical physicist at the University of California, Irvine, uses statistical techniques to analyze the efficacy of immunotherapy for treating triple-negative breast cancer, an aggressive, hormone-independent cancer. From the density and distribution of immune cells in patients, she evaluates the response to immunotherapy and the likelihood of recurrence. "We can predict with about 70% accuracy whether cancer will recur in a patient within five years," she says.

Yu is also interested in why tumors occur where they do. More than half of breast cancers, for example, are found near the armpit. Twice as many lung cancers start in an upper lobe than in a lower lobe. Colon cancer starts preferentially in the first half of the colon. "It's not random. It's a spatial question—a physical property," says Yu. She likes the question "because it's not clear you could answer it by signaling pathways. Something other than signals or toxins is implied."

Princeton's Austin also applies theoretical approaches to approximate and predict cancer behavior. Coupled nonlinear partial differential equations can describe the interactions between cancer cells and the noncancer cells in connective tissue, he says. "It's game theory. You can try to get at what the future will be, how the cancer propagates." Metastasis cannot currently be predicted, he notes. So the value of such calculations would be to estimate outcomes and help prevent overtreatment.

Far more features are visible in cancer tissue than can be identified by eye, says pathologist Roberto Salgado, who works in Belgium and Australia. He integrates machine learning and fractal analysis with spatial genomics, which correlates genetic information in DNA and RNA with location and function of tissue. The shape of cell nuclei, chromatin density, growth patterns of cancer cells, and patterns of blood vessels are some of the variables pathologists use to distinguish between aggressive and more indolent cancers. "Machine learning could complement the work of pathologists," says Salgado. "I think we need to uncover much more of the architectural complexity of cancer, the hidden secrets of cancer morphology. If we do that and then match the spatial architecture of cancer with the genomics, we can make progress for our patients."

One key to meaningful progress in cancer research is improved communication among biologists, clinicians, pathologists, and physical scientists. Last year Memorial Sloan Kettering Cancer Center's Larry Norton and colleagues launched the privately funded Mathematical Oncology Initiative to help build a common language and bring more math to cancer research.

Toni Feder

Questions surround NASA's shutdown of an international cosmic-ray instrument

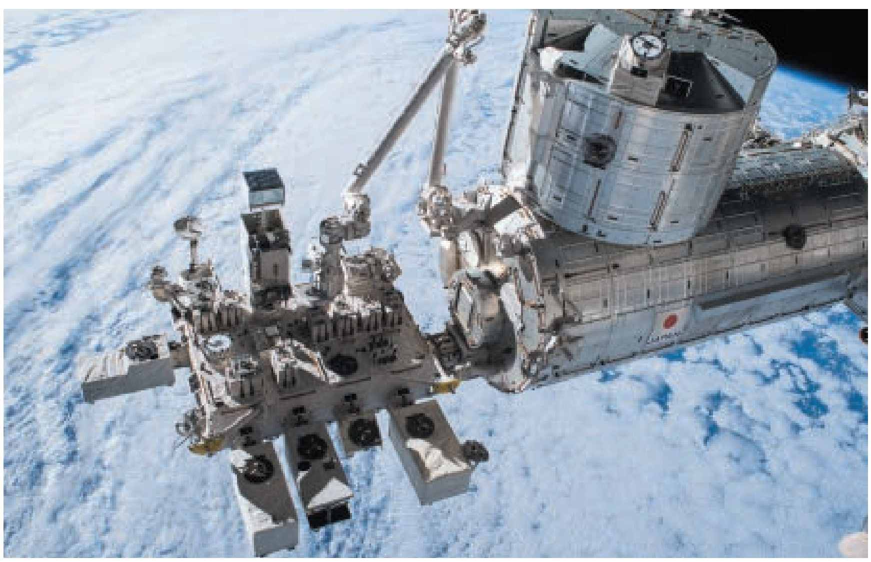
The detector aboard the International Space Station could be turned back on if a new proposal passes peer review. But it's unclear who might operate it.

In February NASA quietly pulled the plug midway through the expected three-year life of a functioning cosmic-ray detector attached to the outside of the International Space Station (ISS). The unusual step came after a majority of scientists in the Cosmic Ray Energetics and Mass for the ISS (ISS-CREAM) collaboration rejected outright NASA's demand to replace the project's principal investigator (PI), University of Maryland (UMD) physicist Eun-Suk Seo, with the agency's hand-picked successor. "We asked the University of Maryland and the science team to make changes," says NASA astrophysics program director Paul Hertz. "They did not make those changes, so we did not continue the mission." A three-sentence note at the top of the ISS-CREAM collaboration's webpage was the only notice of the project's demise. NASA was providing $1.2 million a year for ISS-CREAM operations. It paid $22.4 million to build, launch, and install the refrigerator-sized device on the ISS. South Korea, home to two universities in the collaboration, contributed an additional $10 million.

Hertz says the agency will entertain proposals to resume ISS-CREAM operations. The collaboration "did not generate any science in its first year," he says, noting that reviewers recommended against continuing the project under its existing leadership. "It's up to somebody to write a proposal and demonstrate that if we were to turn it on and give them money, then we would get science, and the science would be worth the money."

ISS-CREAM was installed on the ISS in August 2017 to study properties of high-energy cosmic rays that are believed to originate from the universe's most violent events (see PHYSICS TODAY, May 2010, page 15). The four-instrument detector was adapted for spaceflight from a set of similar instruments, known as CREAM, that were carried aloft on seven high-altitude balloon flights over Antarctica beginning in 2004. Seo was PI on that project.

ISS-CREAM was meant to complement other cosmic-ray detectors, such as the Alpha Magnetic Spectrometer (see PHYSICS TODAY, June 2013, page 12) and Japan's Calorimetric Electron Telescope (both also installed on the ISS), by studying very-high-energy (10^{12} eV to 10^{15} eV) cosmic-ray particles, which range in mass from protons to iron nuclei. ISS-CREAM's two principal instruments are a tungsten sampling calorimeter built by UMD and a silicon charge detector (SCD) contributed by Sungkyunkwan University (SKKU). A third instrument, a boronated scintillator detector (BSD), was built by a team of researchers from the Pennsylvania State University (PSU), Northern Kentucky University, and NASA's Goddard Space Flight Center. Kyungpook National University contributed detectors located above and below the calorimeter to distinguish electrons from protons. France's Laboratory of Subatomic Physics and Cosmology and the National Autonomous University of Mexico also contributed calorimeter components.



NASA

THE ISS-CREAM DETECTOR is the rightmost rectangular shape at the bottom of the photograph. To its right is the Japanese module of the International Space Station.

The ISS-CREAM payload on the space station was expected to generate 10 times as much data as were acquired during the 161 days of balloon experiments, and the cosmic-ray data would be cleaner due to elimination of atmospheric interactions. Before ISS-CREAM, the highest-energy cosmic rays could only be inferred from secondary-particle air showers seen with ground-based telescopes, says Jon Paul Lundquist, a postdoc working under Seo. That indirect process was used to calculate the energy of the highest-energy cosmic ray ever observed (3.2×10^{20} eV), at the now-closed Fly's Eye detector in 1991 (see PHYSICS TODAY, January 1998, page 31).

Internal conflict

A UMD spokesperson declined to comment on the ISS-CREAM cancellation. Seo and other members of her team say that university research administrators had agreed to NASA's demands, which included replacing Seo as PI with Scott Nutter, a Northern Kentucky University physics professor who had been named the collaboration's data manager. "It's natural that the university has to be concerned about its overall relationship with NASA," Seo says. UMD is one of the largest academic recipients of NASA funding.

Since last summer, collaborators from UMD and the South Korean universities had fought with Nutter over the coding of data taken from their respective instruments to a common format for analysis. At a January meeting, UMD and Korean members—representing by far the majority of the collaboration—presented to NASA astrophysics division staff a letter that rejected Nutter's November designation as PI and "disbarred" him from the collaboration. "We didn't say no to NASA. We said no to their choice of PI," says Lundquist.

In the aftermath, NASA turned off the ISS-CREAM device, terminated the UMD grant on one week's notice, and withdrew the grant funds that hadn't been disbursed. Hertz says NASA consulted with the Korean government before the shutdown but notes that the Korean contribution was made to UMD, not through NASA.

Seo says she declined to sign the rejection letter. "I accepted [Nutter] because I had no choice. I had to work with him." She says she was willing to do whatever it took to keep the collaboration alive.

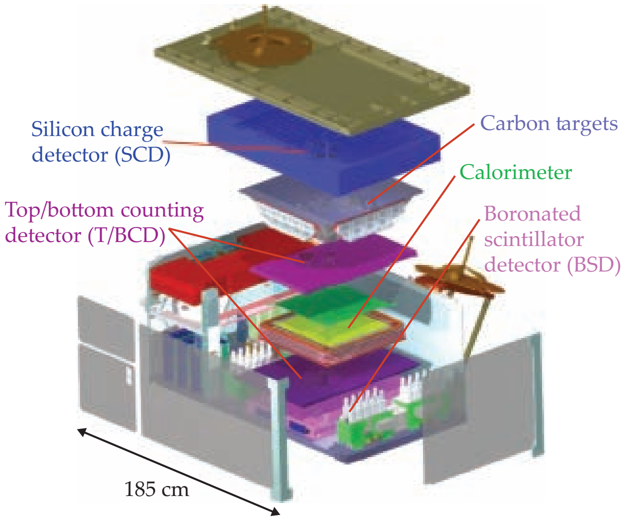
As data manager, Nutter tried to make each detector team organize and structure its data in a way that would smooth collaboration-wide analysis. "He organized a unified software structure we would all have access to, where each instrument team could contribute their own software, but from that archive, all of us could run the full analysis," says PSU physicist Stephane Coutu, a member of the BSD group. "Scott worked hard to bring every part of the instrument together as a whole," he says, including visiting with collaborators in Korea "to help sort out SCD behavior issues." The reviewers too, praised Nutter, saying he had "significantly improved the organization and development of analysis code."

But researchers from UMD and the two Korean teams took an opposing view. Il Park, an SKKU professor who headed the SCD contingent, says Nutter failed to provide collaboration-wide data that incorporated calibrations and computed derived science quantities such as tracking, charge, and energy. Lundquist says the coding changes demanded by Nutter were unnecessary and a distraction from the science. "There was over a decade of code base and data structures that the individuals who know the detectors best created," says Lundquist. Forming a unified structure from that base should have been "extremely straightforward," he says. But Nutter insisted that his coding method be used, Lundquist says.

The Korean collaborators also didn't object to NASA's naming someone other than Nutter as PI, says Park. They suggested Coutu, who declined the job.

Nutter insists that he never wanted the PI job, but he ultimately accepted when NASA officials told him no one else would take it. "The choices were termination or me stepping up to the PI position."

Seo says that NASA would have to clarify the degree of control and oversight the agency would have over a modified mission before she would consider submitting a proposal to resume operating the device. Unusually, NASA did not classify ISS-CREAM as a spaceflight mission, but funded it under a research grant. The project was therefore subjected to a less rigorous set of qualification requirements and agency oversight than are space missions. "We thought [ISS-CREAM] was an interesting idea to try," says Hertz. "Can we take a research-level payload in a research program and put it on the ISS and get space-quality science out of it?"



A SCHEMATIC of the ISS-CREAM instrument showing the four detectors included in the refrigerator-sized device. (Courtesy of the UMD Cosmic Ray Physics Laboratory.)

For ethical and scientific reasons, turning ISS-CREAM over to a different collaboration would be problematic, notes Seo. "It would be a nightmare if anyone else tried to use it and to claim the data, and a nightmare situation for me to accept or not to accept results" if they conflicted with those of the original collaboration.

Disbarring NASA's choice

In January 2018, six months into the mission, a NASA review team identified problems with ISS-CREAM, including an understaffed instrument control center at UMD, lagging data analysis, and poor communications among the various instrument teams. Based on the review recommendations, in April NASA appointed Nutter as data manager and brought in former Goddard engineer James Dickey to a new position of mission operations manager.

A follow-up in September 2018 by NASA and external reviewers reported that both data management and mission operations had improved significantly. But the review noted that only 50 days' worth of science-quality data had been collected during ISS-CREAM's first year and said competent leadership of the calorimeter measurements and analysis—the UMD team's responsibility—was lacking.

Without changes to the project's leadership, the review stated, "it is not likely that significant scientific results will be produced within the remaining time of the nominal three-year mission." It recommended replacing Seo as PI with "someone with sufficient knowledge and experience in cosmic ray science, familiarity with the details of the four subsystems, and clear leadership capability." In the event a new PI couldn't be found, the review recommended termination, "as additional investments are unlikely to result in substantial scientific return under the current leadership of the project."

Seo and other collaborators take issue with Hertz's and reviewers' complaints regarding ISS-CREAM's low scientific productivity. Researchers were on track to meet the grant application's timetable for reporting results, says Lundquist, who notes that 10 abstracts for presentations of results have been submitted for the International Cosmic Ray Conference next month.

As for the relative dearth of usable data, Park says that it took the better part of a year to calibrate the instrument and be assured that it wouldn't be damaged by energetic particles during ISS transits of the South Atlantic Anomaly in Earth's inner Van Allen radiation belt. Such a lengthy break-in period is "typical and always true for all high-energy particle experiments," he says.

Out of work

Five postdocs and one other scientist on the UMD team, meanwhile, have continued to analyze data from the terminated mission without pay. Lundquist says he is moving to South Korea to work at SKKU for a couple months at half the salary he'd been getting from the NASA grant.

NASA offered the six out-of-work UMD scientists positions at the nearby Goddard campus to work on closing out ISS-CREAM. It also offered to consider them for unspecified permanent positions. But in a letter to NASA's astrophysics division manager, all six declined, saying the offer appeared to be an effort to further divide the ISS-CREAM team.

Lundquist says the episode should raise concerns about the degree to which NASA micromanages its grantees. "It seems highly unusual for a funding agency to reach their hands into a collaboration and demand these specific conditions," he says.

But Hertz says that NASA conducts reviews of all its space missions at the end of their "prime mission," period. "I would say that termination happens seldom but not never."

Other astrophysics missions were terminated after such reviews were performed, including the *Wide-field Infrared Survey Explorer* and the *Extreme Ultraviolet Explorer*. *WISE* was put into hibernation in 2011 after its hydrogen coolant was depleted, but it was later reactivated to look for near-Earth asteroids (see PHYSICS TODAY, March 2015, page 19). *EUVE* operated for more than eight years before its deactivation in 2001.

As a research grant, ISS-CREAM's prime mission period was one year, with extended operation contingent on passing reviews, Hertz says.

David Kramer PT

Applied Physics Reviews

5 FACTS ABOUT APR

1 High Impact Factor

2017 Impact Factor: **12.894**

Put your research in front of the right audience
With an IF of **12.894**, APR is an ideal journal for articles on important topics in experimental or theoretical research in applied physics or applications of physics to other branches of science.

2 Cutting Edge Discoveries

Accepts Original Research Articles and Reviews
Historically a review journal, APR began considering submissions of original research for publication in January 2019. Future research articles will report on important and novel research studies of general interest to the applied physics community.

3 Invested in the Future of Science

APR is published by a Not-for-Profit Organization
Revenue from the journal helps support the scientific and educational programs of the American Institute of Physics such as The Society of Physics Students, Neils Bohr Library & Archives and InsideScience.org.

4 Superior Author Experience

Publishing should be easy, not stressful
APR's Editorial Team is here for you throughout the submission, review, and publication process and we'll treat your manuscript like it's our own. You will receive expert editorial guidance and fair and constructive peer review. Additionally our Production Team will provide expertise in copy editing to ensure clarity and grammatical correctness, making your paper the best it can be.

5 Awesome Editorial Team

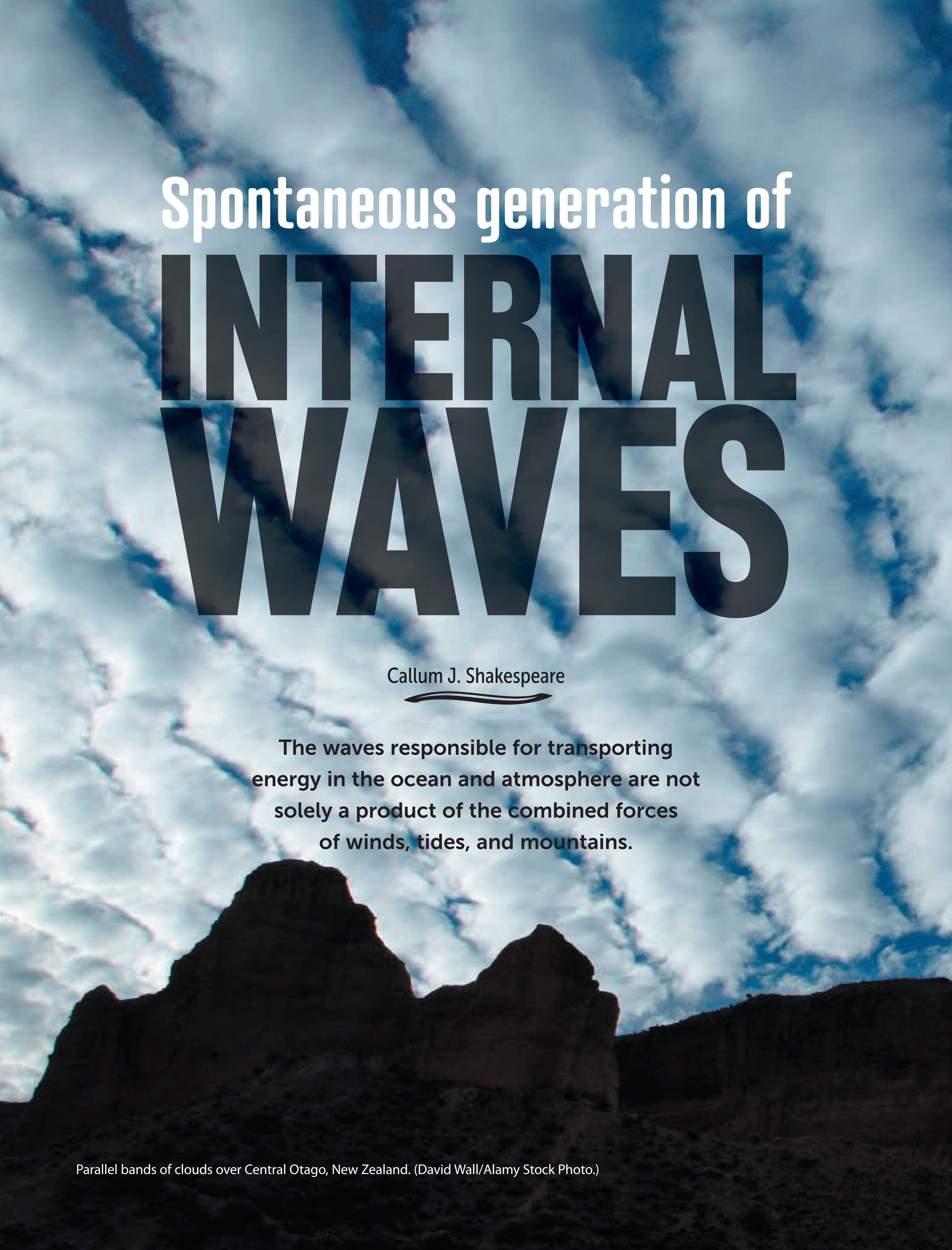
Receive professional and expert feedback
Luigi (pictured) is real! He is one of several Ph.D. scientists on our Editorial Team. They each have active experience in reviewing and publishing papers and their combined experience provides the foundation for the premium service and attention your research will receive. Read about Luigi and the rest of the team on our website **products.aip.org/apr**

Submit today!

Meet Luigi and the rest of the APR Editorial Team. Watch our video!

Visit apr.aip.org/video

AIP Publishing



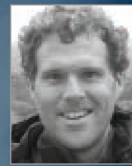
Spontaneous generation of INTERNAL WAVES

Callum J. Shakespeare

The waves responsible for transporting energy in the ocean and atmosphere are not solely a product of the combined forces of winds, tides, and mountains.

Parallel bands of clouds over Central Otago, New Zealand. (David Wall/Alamy Stock Photo.)

Callum Shakespeare is a research fellow at the Australian National University in Canberra.



Look up at the sky on a cloudy day, and you often see sets of parallel, equally spaced bands of clouds that are the signature of a type of wave known as an internal wave. Although visible under only the right conditions, internal waves are ubiquitous in both the atmosphere and the ocean. In many dynamical systems, waves release excess energy in a fluid that is displaced from its lowest-energy, balanced state. Internal waves extract and transport energy three-dimensionally before releasing it to large-scale circulation.

Historically, researchers thought that only mechanical forcing or direct thermal displacement of the system from its balanced state could generate internal waves. However, more recent work shows that a spontaneous imbalance of the fluid system can generate or amplify them without any direct forcing. The discovery has led to new perspectives on the role of waves in the circulation of the atmosphere and ocean. (For more on internal waves in planetary atmospheres, see the article by Erdal Yiğit and Alexander S. Medvedev on page 40 of this issue.)

Waves that you see on the surface of an ocean are near cousins of internal waves, and fluid dynamicists call them surface gravity, or just surface, waves. Strong storms generate surface waves, which are periodic oscillations in the height of the water, and they can propagate 10 000 km across entire ocean basins before they break on your local beach. The motion of surface waves is confined to the two-dimensional surface of the ocean. In contrast, internal waves, which storms also generate, propagate three-dimensionally through the ocean's interior as periodic oscillations in the height of stratified constant-density layers. For some internal waves, the oscillations are a hundred meters or more in amplitude—far larger than surface waves.

All waves have a dispersion relation, which quantifies the relationship between wavelength and frequency in terms of certain physical parameters. The two important parameters for internal waves are the Coriolis frequency and the buoyancy frequency. The Coriolis frequency f is the vertical component of Earth's rotation vector at a fixed point on its surface and depends on latitude. The value of f sets the minimum frequency for internal waves. Those oscillating at that frequency are called inertial waves, and they propagate horizontally, as shown in figure 1.

The buoyancy frequency N is proportional to the vertical rate of change in the density of the atmosphere or ocean. It is the natural frequency of oscillation when a volume of dense fluid is displaced into the lighter fluid above or vice versa, and it determines the maximum frequency for internal waves. Those at the buoyancy frequency are called gravity waves and propagate purely in the vertical. Internal waves with intermediate frequencies propagate both horizontally and vertically.

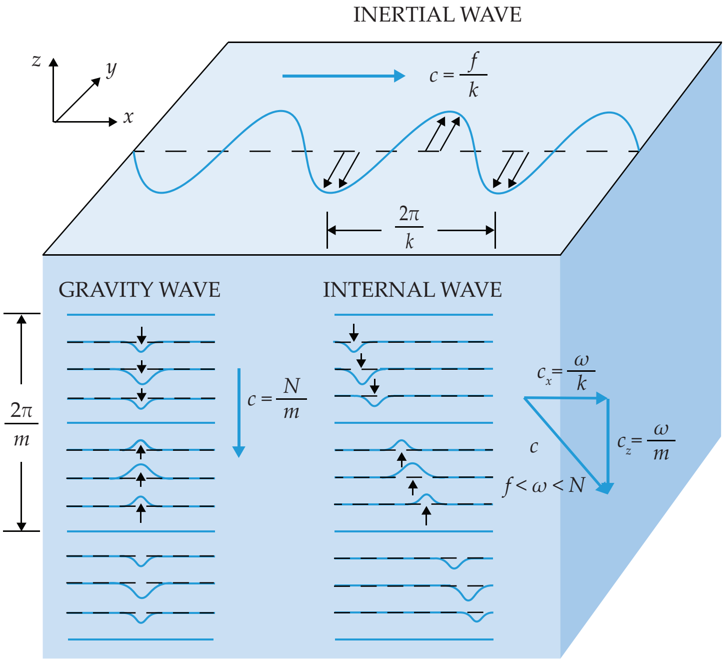


FIGURE 1. INTERNAL WAVES. The lowest-frequency internal waves are inertial waves, which oscillate at the Coriolis frequency f given by the vertical component of Earth's rotation vector at a fixed point on its surface. They propagate horizontally at speed $c = f/k$, where k is the horizontal wavenumber. The highest-frequency internal waves, gravity waves, oscillate at the buoyancy frequency N and propagate vertically at speed $c = N/m$, where m is the vertical wavenumber. Intermediate-frequency internal waves propagate both horizontally and vertically.

In the midlatitudes, internal waves in the atmosphere and in the ocean are typically high frequency, with periods of 20 hours or less. By contrast, most large-scale circulation in the ocean and atmosphere varies slowly enough—with periods of months to years—that it may be treated as balanced motion.[1] In geophysical fluid dynamics terms, balanced motion does not accelerate relative to a fixed point on Earth's surface. In the atmosphere and ocean, much of the circulation is in thermal-wind balance, where the pressure resulting from a horizontal density gradient is counteracted by the Coriolis force from Earth's rotation. Such density gradients, or fronts, in thermal-wind balance are abundant in the atmosphere and ocean. Atmospheric high- and low-pressure systems—and their ocean equivalent, known as mesoscale eddies—also exist in a state of near thermal-wind balance. However, if all those systems are balanced, how do they lose energy, decay, and otherwise evolve over time? And how are unbalanced motions such as waves generated?

Internal waves transport significant energy and momentum from their sources, usually near fluid boundaries, into the interior of both the atmosphere and ocean. If the waves have large enough amplitude and small enough scale, they break and dissipate, causing mixing and the acceleration of the balanced flow. A wave's growth is a consequence of many different mechanisms, including changes to the density structure of the propagating medium, interactions with currents, constructive interference with other waves, and other interactions. In the atmosphere, breaking internal waves contribute to the poleward flow of air in the middle atmosphere (10 km to 80 km high) and therefore help sustain the Brewer–Dobson circulation, wherein air originating at the surface in the tropics cycles to higher altitudes then poleward. Thus, internal waves are a crucial component of weather and climate models.[2] Wave breaking is also vitally important in the ocean abyss, below a depth of 2 km, where it drives the mixing of dense water with the lighter waters above (see the article by Adele Morrison, Thomas Frölicher, and Jorge Sarmiento, PHYSICS TODAY, January 2015, page 27). Without that mixing, there would be no overturning of the deep ocean.

Forced generation of internal waves

Have you ever been on a commercial airliner flying over a mountainous region and been warned to return to your seat and buckle up? The sky outside the windows may even have looked clear, but the aircraft soon started shaking and vibrating. That motion was due to a region of clear-air turbulence typically caused by the breaking of a mountain wave.[3] Water or air flow over surface topography, or orography, generates mountain waves, also known as orographic waves; they are the most studied and most prevalent type of internal wave in the atmosphere. Some mountain waves break at low levels near the mountains. Others break in the lower stratosphere, at the typical cruising altitudes of jet airliners, and present a significant aviation hazard.

Mountain-wave generation relies on

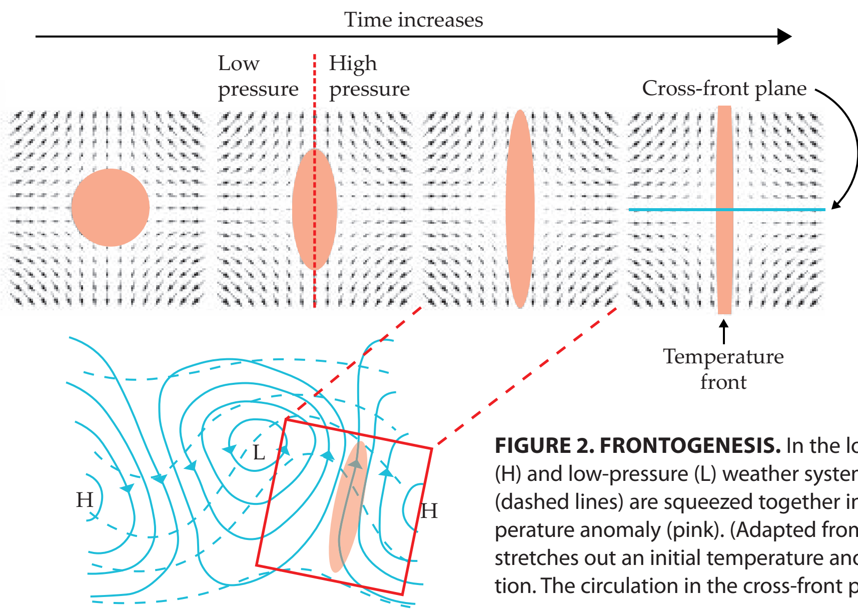


FIGURE 2. FRONTOGENESIS. In the lower image, a front forms between high-pressure (H) and low-pressure (L) weather systems; isobars are shown as solid lines. Isotherms (dashed lines) are squeezed together in regions of confluent flow, where there is a temperature anomaly (pink). (Adapted from ref. 5.) In the upper image, the flow (vectors) stretches out an initial temperature anomaly (pink circle, left) into a front through advection. The circulation in the cross-front plane (blue line) is shown in figure 3.

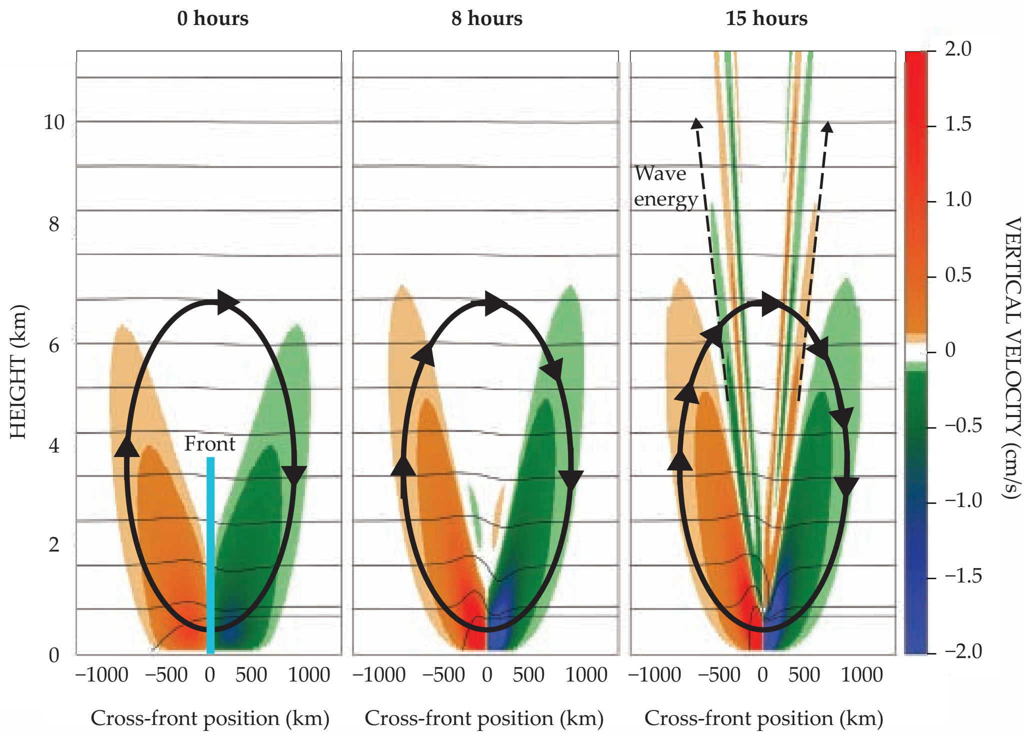


FIGURE 3. FRONT FORMATION AND SPONTANEOUS GENERATION of internal waves with parameter values appropriate for a weather front. Initially (left), the front is smooth with a broad and weak upwelling on the warm side and downwelling on the cool side. The thin black lines are isotherms, and the thick black arrows indicate the circulation. Eight hours later (center), the front has sharpened and exhibits stronger circulation. After 15 hours (right), the front has sharpened to a point where spontaneous generation is possible. The generated waves, indicated by dashed arrows, propagate up and away from the front.

sufficiently fast near-ground flows and a stably stratified atmosphere, which provides the restoring force on the wave. The topographic obstacle, of length L, pushes the flow, at speed U, up against the force of gravity and over it. If the rate of deflection exceeds the minimum frequency of an internal wave—that is, if $U/L > f$—then waves are radiated. The nondimensional parameter $U/(fL)$ is known as the Rossby number, Ro, after famed fluid dynamicist Carl-Gustaf Rossby (see the article by Jim Fleming, PHYSICS TODAY, January 2017, page 50), and is one of the most important parameters in meteorology and oceanography.

Mountain waves also occur in the deep ocean, when abyssal currents flow over seafloor topography. However, the currents are slow, with a smaller Rossby number, compared with their atmospheric counterparts. As a result, mountain waves are less common in the ocean than in the atmosphere. Instead, the predominant mechanism for wave generation in the deep ocean is the interaction of the tide with seafloor topography. The gravitational pull of the Moon—and to a lesser extent, the Sun—drives the tide, which is a daily or higher-frequency bulk motion of the water column. That sloshing of stratified water back and forth over the seafloor directly radiates internal waves at the tidal frequency. Tidal generation contributes about 1.5 TW of energy to the ocean wave field; mountain waves[4] add only 0.2 TW. A final source of internal waves unique to the ocean has already been mentioned—storms or, more precisely, air currents periodically forcing the ocean surface from above. Storms contribute about 0.3 TW of energy to the ocean's internal wave field.

The processes described thus far are examples of forced generation of internal waves. They rely on direct high-frequency forcing, such as tides or winds, or interaction with an external body—for example, mountains. Let us turn our attention to the generation of waves without forcing or interaction.

Frontogenesis and spontaneous wave generation

The theory of spontaneous generation originated in the study of weather fronts and their formation, or frontogenesis. Because the fronts are associated with extreme weather such as rain, hail, and snow, frontogenesis has been the focus of much discussion in atmospheric literature over the past century. It occurs in the region between high- and low-pressure systems where the flow is converging, as shown in figure 2. Such flow convergence squeezes and sharpens any existing anomalously steep temperature gradient into an identifiable temperature front.

In 1968, during his PhD work at the University of Cambridge, Brian Hoskins formulated a mathematical model of frontogenesis.[5] It describes how a circulation develops around a front that is initially in thermal-wind balance. The confluence of the high- and low-pressure systems drives circulation with fluid moving up on the warm side and down on the cool side. Figure 3 shows the predicted circulation and temperature structure across the front over time. In the figure, the circulation intensifies as the front sharpens over eight hours. A central tenet of the Hoskins model is that the front remains in a balanced state throughout its evolution. As it sharpens, however, that no-acceleration assumption begins to break down, and internal waves form spontaneously.

The mathematical description of spontaneous generation requires an extension of the Hoskins model to include the unbalanced cross-frontal acceleration that leads to waves. In a case of academic symmetry, I developed that extended model in 2014 while working on my PhD at the University of Cambridge.[6] The extended model departs from the Hoskins model only when the front becomes sufficiently sharp and the convergence of hot and cold regions is sufficiently strong.

Mechanistically, spontaneous internal waves are associated with the rapid vertical displacement of fluid in an otherwise stably stratified environment, similar to mountain waves. Consider a small amount of fluid, a "parcel," on the warm side of the front. The large-scale convergence pushes the parcel toward the front and then upward, against the restoring force of gravity, into the cooler ambient air above. If that process occurs faster than the minimum frequency of an internal wave, then a wave is generated.

As the front sharpens, the circulation increases, as shown in

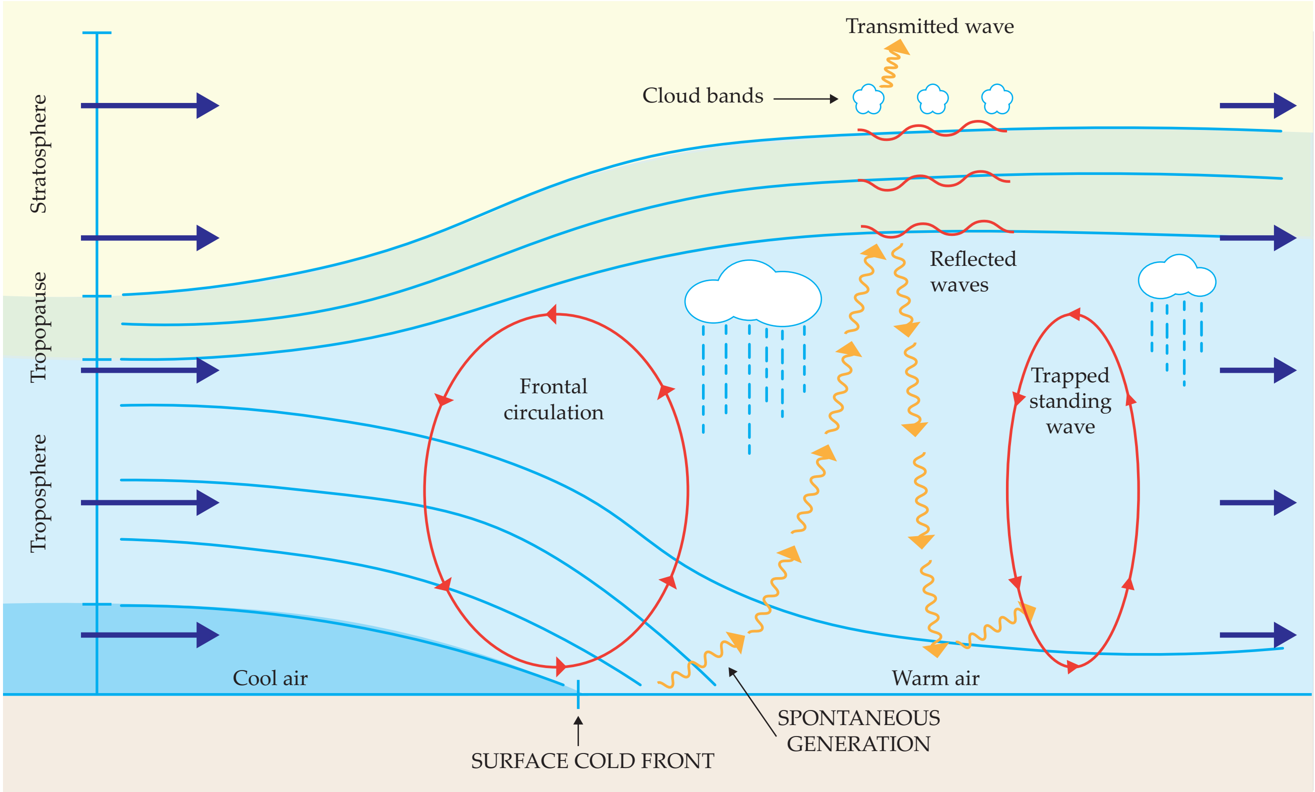


FIGURE 4. AN ATMOSPHERIC COLD FRONT with large-scale confluent flow moving left to right, which drives circulation (red arrows) around the front in the plane of the page. The rising air ahead of the front causes clouds and precipitation. Waves are spontaneously generated at the surface front and propagate ahead of it (yellow arrows). Some propagate through the tropopause into the stratosphere; they create horizontal bands of clouds at the interface through the introduction of moist tropospheric air. Ultimately, those waves transfer momentum to the stratospheric circulation. Other waves reflect and form troposphere-trapped standing waves, which may themselves cause clouds and rain.

figure 3. Oncoming convergent flow deflects faster around the front and will reach the minimum-frequency criterion unless frontogenesis is halted by frictional processes or dynamical instabilities. A sharper front leads to a higher amplitude and smaller scale of the generated wave. The convergent flow traps the wave in a location where its horizontal phase speed is equal and opposite to the flow speed, such that the wave appears as a nearly steady feature of the flow. However, the wave propagates vertically and carries energy upward. The idealized mathematical model predicts symmetric wave generation on either side of the front, but in practice, generation occurs preferentially on the warm side, where the surface gradients are larger.

Waves generated spontaneously at atmospheric fronts have a significant effect on the weather. A cold front, shown moving to the right in figure 4, sharpens due to large-scale convergent flow and generates frontal circulation. The rising air ahead of the front forms clouds. When the front sharpens sufficiently, it generates waves (yellow wavy arrows) propagating upward and ahead of the front. Some generated waves, typically smaller-scale ones, propagate directly into the stratosphere. They are responsible for the long, parallel bands of clouds seen on page 34, and they contribute to the large-scale atmospheric circulation when they break and transfer momentum to the balanced flow.

Other generated waves, typically larger-scale ones, contribute to the formation of storms ahead of the advancing cold front.[7] The formation relies on a weakly stratified region in the upper troposphere that reflects upward-propagating waves, which then interfere with the reflected waves. A standing wave builds and propagates ahead of the front until it becomes stuck in the oncoming convergent flow. The wave stays ahead of the front and creates one or more narrow bands of vertical motion with typical widths of one to tens of kilometers. Those bands produce clouds, precipitation, and strong winds.[8]

Frontogenesis and spontaneous generation are not unique to the types of flow described thus far, although those are the most common frontogenetic scenarios. Researchers observe similar effects in fronts with flows whose velocity changes with height.[9] All mechanisms of spontaneous generation share a common property: The amplitude A of the generated wave is exponentially small[1,6] in Rossby number, $A \sim e^{-1/Ro}$. Physically, the result means that spontaneous generation exhibits a threshold behavior: Amplitudes are significant only once the Rossby number exceeds a critical value, which theory estimates to be 0.15–0.20. An immediate consequence is that spontaneous generation is localized in time and space to regions that satisfy the threshold.

The ocean: Spontaneous or stimulated?

One flow regime in which the Rossby number threshold is routinely exceeded is the ocean submesoscale, a scatter of density fronts and eddies with lateral scale of less than 10 km and lo-

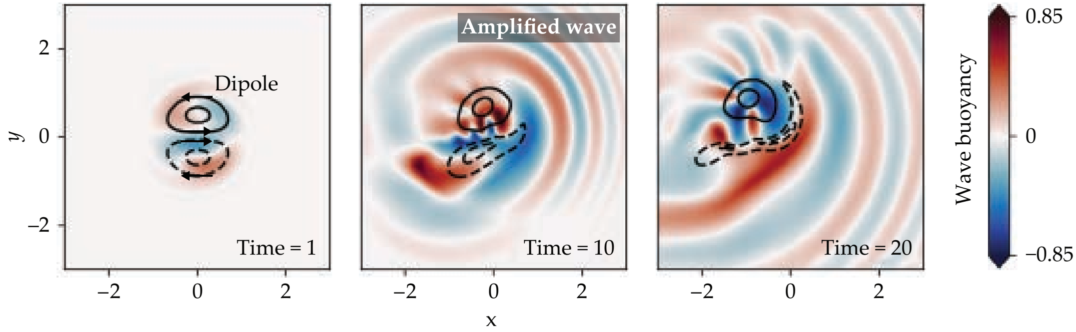


FIGURE 5. STIMULATED EMISSION occurs when a preexisting internal wave (color) interacts with a dipole (black contours show flow), or double eddy. In the panel sequence, the dipole weakens with time as its energy is transferred to the wave. The wave's height, given here in terms of wave buoyancy, increases with time as it gains energy. All quantities are expressed in nondimensional units. (Adapted from ref. 13.)

calized primarily in the upper 50 m of the ocean.[10] Large strains and shears from the surrounding turbulent flow prime the density fronts at the submesoscale for spontaneous generation. Recent numerical models attempted to quantify the global energy transfer to waves associated with spontaneous generation in that region.[11,12] Those studies found that although local energy flux from spontaneously generated waves may be large, the area-averaged values are relatively small and account for 0.03 TW of wave generation globally, or 1/10 of the 0.3 TW due to winds. Spontaneous generation contributes to, but is not a dominant source of, wave energy in the ocean.

Spontaneous generation is one mechanism for transferring energy from balanced flow to unbalanced waves. The search for other links between balanced and unbalanced flow is a major focus of the oceanography community. Recent work proposes a new mechanism called stimulated emission, the transfer of energy from balanced flow to a preexisting wave rather than the generation of new waves. Stimulated emission does not require a large Rossby number and therefore transfers much more energy. Theorists have developed models to describe the dynamics of stimulated emission at low Rossby number, where the balanced flow is well understood.[13,14] The results of one model are shown in figure 5. The model starts with a uniform wave and a localized dipole, or double eddy. The dipole locally squeezes and amplifies the gradients of the wave, and the wave's energy increases with time as the dipole loses energy. Studies using more realistic models[15,16] suggest that stimulated emission occurs frequently in the ocean, especially in response to wind forcing. One study[15] reported a 30% enhancement, or about 0.1 TW globally, in the energy in wind-generated internal waves, at the expense of balanced flow.

Outlook

Many questions remain about the importance and role of spontaneous generation and related processes, such as stimulated emission, in creating internal waves. Those questions have important consequences for the fields of oceanography, meteorology, and climate science.

In the ocean context, recent work on spontaneous and stimulated emission showed that internal waves sometimes gain energy from and lose energy to balanced flows.[12] That body of work shifts the field away from the classical paradigm that winds, tides, and topography force internal waves, which then directly dissipate their energy and drive mixing. The energy pathways between generation and dissipation are considerably more complex, and understanding them will lead to better constraints on the variability and location of wave-driven mixing and the associated effects on ocean circulation.

In the atmosphere context, researchers are interested in quantifying nonorographic sources of internal waves. Midlatitude fronts are major sources of nonorographic waves, and spontaneous generation is one mechanism for generating those waves. Although nonorographic waves are locally less intense than mountain waves, which are highly localized with large amplitudes, their cumulative contribution may be comparable.[17] They are often not included in global atmospheric models because of the uncertainty in their spatial and temporal distribution. Theoretical advances are starting to correct the situation, but more work is needed to translate idealized theoretical models into more realistic settings. Given that nonorographic waves may contribute substantially to circulation in the middle atmosphere, their inclusion in models should improve seasonal and multiyear forecasting.

Researchers have made significant progress in their understanding of internal waves in the ocean and atmosphere through a combination of observational, modeling, and theoretical advances, only a small subset of which are discussed here. Spontaneous generation is one element of that work, which broadly examines the interactions of unbalanced flows with the large-scale balanced flows that dominate our climate and weather.

REFERENCES

1. J. Vanneste, *Annu. Rev. Fluid Mech.* **45**, 147 (2013).
2. D. C. Fritts, M. J. Alexander, *Rev. Geophys.* **41**, 1003 (2003).
3. A. D. Elvidge et al., *Meteorol. Appl.* **24**, 540 (2017).
4. A. F. Waterhouse et al., *J. Phys. Oceanogr.* **44**, 1854 (2014).
5. B. J. Hoskins, F. P. Bretherton, *J. Atmos. Sci.* **29**, 11 (1972).
6. C. J. Shakespeare, J. R. Taylor, *J. Fluid Mech.* **757**, 817 (2014).
7. D. M. Schultz, *Mon. Weather Rev.* **133**, 2449 (2005).
8. C. J. Shakespeare, *J. Atmos. Sci.* **73**, 2837 (2016).
9. F. Lott, R. Plougonven, J. Vanneste, *J. Atmos. Sci.* **69**, 2134 (2012).
10. J. C. McWilliams, *Proc. R. Soc. A* **472**, 20160117 (2016).
11. T. Nagai et al., *J. Phys. Oceanogr.* **45**, 2381 (2015).
12. C. J. Shakespeare, A. M. Hogg, *J. Phys. Oceanogr.* **48**, 343 (2018).
13. C. B. Rocha, G. L. Wagner, W. R. Young, *J. Fluid Mech.* **847**, 417 (2018).
14. J.-H. Xie, J. Vanneste, *J. Fluid Mech.* **774**, 143 (2015).
15. R. Barkan, K. B. Winters, J. C. McWilliams, *J. Phys. Oceanogr.* **47**, 181 (2017).
16. S. Taylor, D. Straub, *J. Phys. Oceanogr.* **46**, 79 (2016).
17. R. Plougonven, F. Zhang, *Rev. Geophys.* **52**, 33 (2014).

PT

Erdal Yiğit is an associate professor of physics at George Mason University in Fairfax, Virginia. **Alexander S. Medvedev** is a researcher at the Max Planck Institute for Solar System Research in Göttingen, Germany.

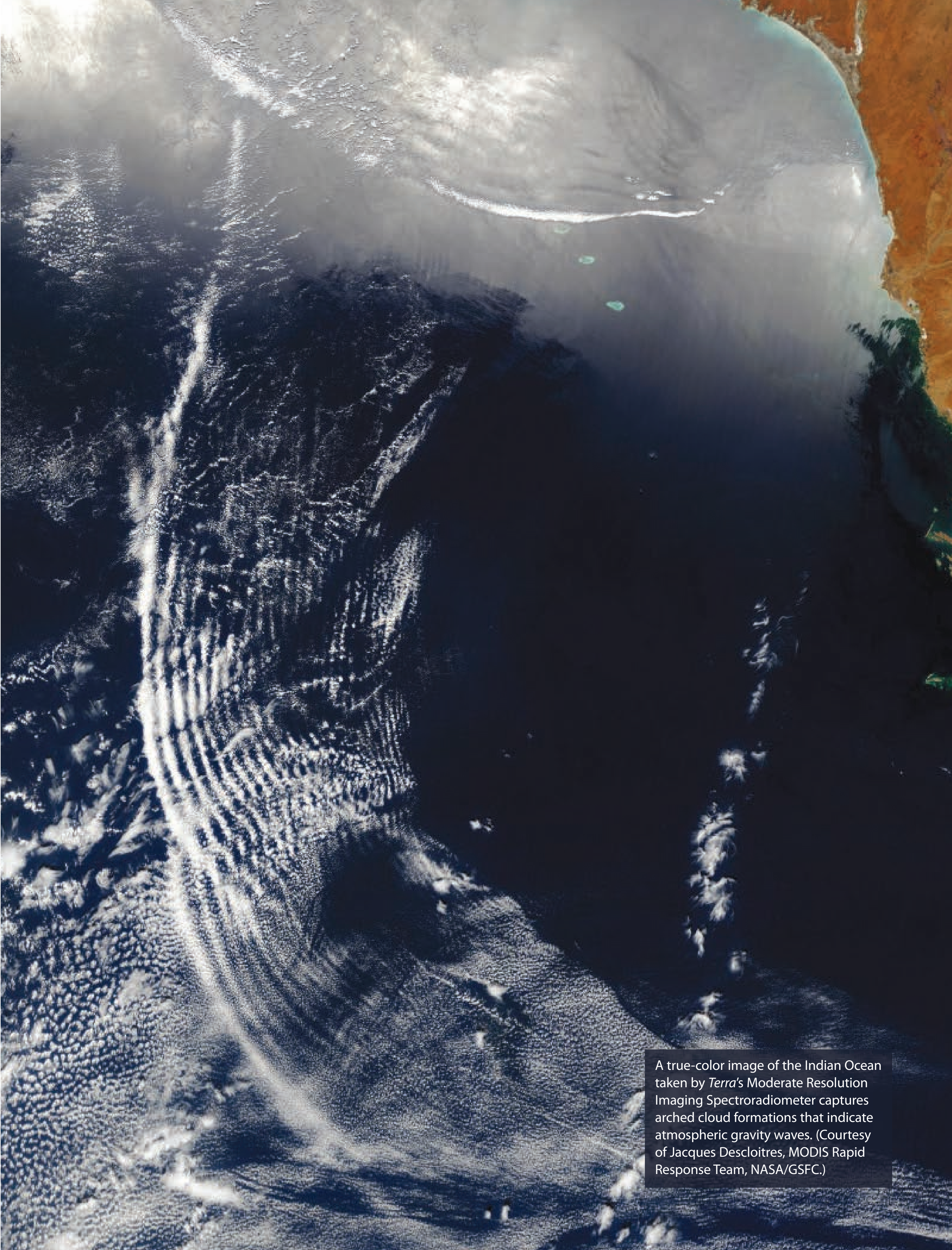
Obscure waves in planetary atmospheres

Erdal Yiğit and Alexander S. Medvedev

On Earth and on other planets, internal gravity waves shape the dynamics and thermodynamics of the atmosphere.

In 1893 the Arctic explorer Fridtjof Nansen experienced the phenomenon of dead water when some mysterious force brought his ship to an almost complete stop. Seafarers had long known that effect and superstitiously linked it to cursed drowned sailors holding onto the vessel. Nansen later asked a doctoral student, Vagn Walfrid Ekman, to explain the phenomenon scientifically. Ekman's thesis became the first academic work dedicated to explaining internal gravity waves. He attributed the dead-water phenomenon to lighter fresh water from melting sea ice sitting on top of saltier, denser seawater. The slowly moving vessel induced waves at the interface between the two layers beneath the sea surface, so its propulsion energy was wasted in making those waves instead of moving the ship. (For more on internal waves in the ocean, see the article by Callum Shakespeare on page 34 of this issue.)

Waves are ubiquitous in both the atmosphere and the ocean. The most familiar are sea waves: Almost everyone has observed how water waves travel along the ocean's surface, grow in size, and then break as they approach the coast. Waves can be viewed as a collective phenomenon resulting from oscillations around equilibrium of many neighboring parcels of a material. They can transfer energy without any net transport of material. Gravity waves, like those considered by Ekman at the interface between two types of water, are similar to those on the sea surface that divide dense water and tenuous air. When a fluid parcel is displaced vertically, the forces that counteract the displacement are buoyancy and gravity. Consequently, such waves are alternatively called buoyancy or gravity waves. (The latter should not be confused



A true-color image of the Indian Ocean taken by *Terra*'s Moderate Resolution Imaging Spectroradiometer captures arched cloud formations that indicate atmospheric gravity waves. (Courtesy of Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.)

with gravitational waves from the theory of general relativity as studied in cosmology.)

In a fluid with continuously varying density, oscillating parcels successively push the fluid above and below them during each cycle, so waves can propagate vertically in addition to cruising horizontally along the surface of constant density. To emphasize the distinction between surface and internal waves, those that can propagate both vertically and horizontally are called internal waves.

The atmosphere is a fluid with continuously varying density, so it is subject to the gravity waves described above. Although they usually get very little attention, they are omnipresent and important. Figure 1, a snapshot from a computer simulation conducted with a general circulation model, shows the signatures of gravity waves in the atmosphere at an altitude of around 250 km.

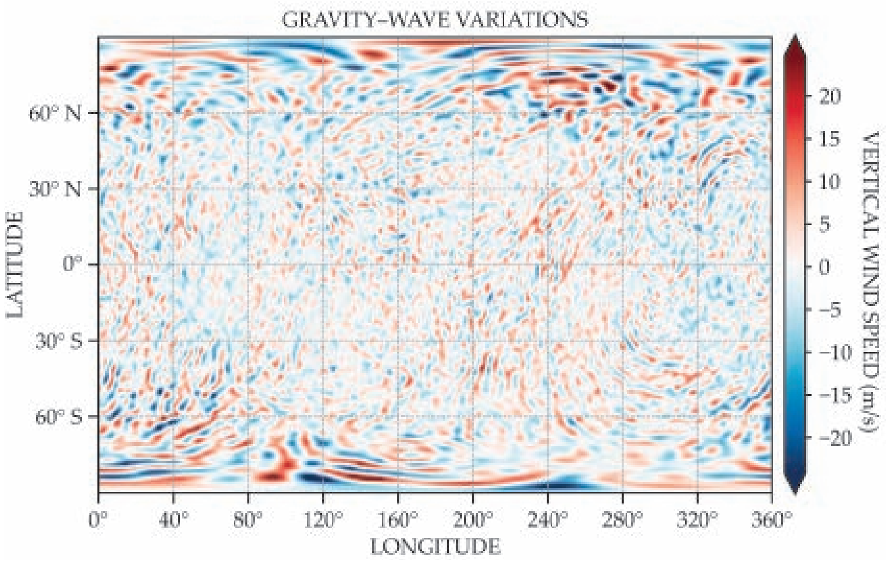


FIGURE 1. THE VERTICAL COMPONENT OF THE NEUTRAL WIND at an altitude of 250 km, simulated with the ground-to-topside model of atmosphere and ionosphere for aeronomy, is a proxy for gravity waves.[16] This snapshot shows the latitude–longitude distribution of vertical wind speed on 3 January 2009, 00 UT. A horizontal resolution of 1° x 1° allows the simulation to reproduce harmonics with horizontal wavelengths of 375 km or greater. (Courtesy of Yoshizumi Miyoshi.)

Gravity waves in the atmosphere

Air parcels rarely oscillate purely in the vertical direction. They have horizontal displacements as well, so their orbits are tilted ellipses. The more inclined the ellipse is, the longer it takes for a parcel to return to its initial state. Thus the periods of gravity waves are tightly linked to their horizontal and vertical wavelengths,[1] which range from a few minutes to many hours but are less than a day. The horizontal length scales of gravity waves extend from a few tens of kilometers to many hundreds but remain significantly smaller than the planetary radius.

Gravity waves represent only a part of the spectrum of vertically propagating internal waves. Planetary waves that define weather systems span thousands of kilometers and last for days, which makes gravity waves appear to be small-scale and have short periods. On the other hand, gravity waves are very large and low frequency when compared with sound waves, a familiar class of atmospheric oscillations.[2]

Earth's atmosphere can be divided into distinct layers based on its temperature profile, as depicted in figure 2. Internal gravity waves are continuously produced by various processes in all layers, but they primarily originate in the lower atmosphere, or troposphere. Since vertical displacements of air are the necessary condition for their excitation, any process that shifts parcels vertically can potentially generate an internal wave.[3] Some obvious although infrequent sources of such waves in the atmosphere are tsunamis, hurricanes, earthquakes, and volcano eruptions. Air flows over hills and mountains are a more common generation mechanism. However, the main source of gravity waves in the troposphere is weather. Air parcels are forced to move vertically during such meteorological processes as convection, atmospheric fronts, cyclonic activity, and instability in wind systems. Because they have so many sources, gravity waves have phase velocities ranging from zero to an upper limit approaching the speed of sound in the atmosphere.

Gravity waves carry both energy and momentum. The energy of an individual wave or wavepacket is proportional to the product of its amplitude squared and the average density of the ambient air. As the wave propagates upward, its amplitude grows exponentially to compensate for the atmosphere's steeply declining density. That process changes the picture of motion in the middle and upper atmospheres: Gravity waves can't be ignored because they actually dominate. The middle and upper atmospheres are like a stormy ocean rippled with huge velocity disturbances accompanied by temperature fluctuations of tens of degrees, which makes the occasionally disastrous weather on Earth's surface seem like a relatively calm seafloor.

Historical retrospective

Atmospheric gravity waves were largely ignored until the golden age of aviation in the 1920s and 1930s, when understanding them became of practical interest. Aircraft occasionally encounter bumpiness when flying over hills and mountains, and often that bumpiness is related to gravity waves generated by that topography or to turbulence from breaking waves. Sometimes the bumpiness is accompanied by visual cues in the form of parallel bands of clouds left behind by propagating wavepackets. However, the most dangerous situation is clear-air turbulence (CAT), which, as the name suggests, is not manifested by clouds or linked to terrain features; its relation to weather phenomena is unclear. Understanding and predicting CAT is therefore important so that pilots can be prepared to deal with it. The dominant idea today is that CAT is the pseudoturbulence created by a superposition of many gravity waves,[4] but forecasting it remains a challenge for scientists and the aviation industry.

Meteorologists developed a theoretical description for grav-

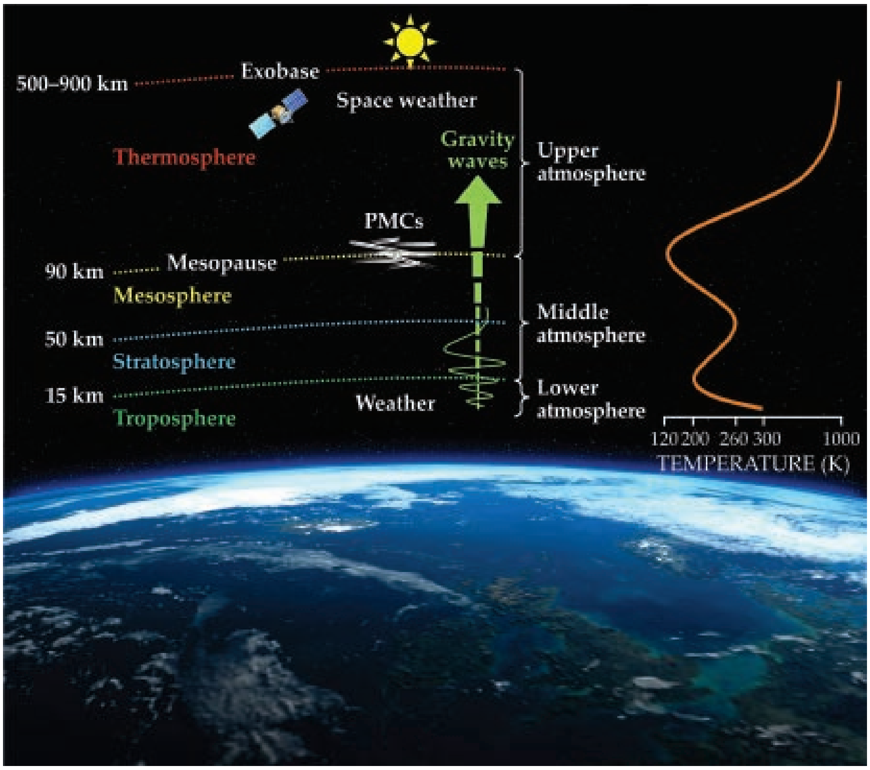


FIGURE 2. THE VERTICAL STRUCTURE of Earth's neutral atmosphere comprises layers that are identified by their temperature (orange line). Note that the temperature scale shown is nonlinear. The atmosphere's extent can be from 500 km to 900 km, depending on external forcing such as solar and geomagnetic activity. Weather processes and the generation of most energetic gravity waves take place in the troposphere, which is the lowest layer and extends about 10–15 km above Earth's surface. The middle atmosphere, containing the stratosphere and mesosphere, is above the troposphere and is separated from the thermosphere by the mesopause at an altitude of about 90 km. Polar mesospheric clouds (PMCs) that trace propagating gravity waves occur in the mesosphere. The upper atmosphere generally refers to the tenuous and hot thermosphere. Space weather processes—the effects of the Sun on the geospace environment—primarily influence the upper atmosphere.

ity waves in the early part of the 20th century[5] but then lost interest in them, mainly because they were seen as weak and unimportant for weather. The energy associated with gravity-wave disturbances is, on average, only a small percentage of what is found in larger, smoother motions in the troposphere. Moreover, the waves were seen as a nuisance in the early days of numerical weather predictions because including them in simulations was computationally expensive, especially considering the computing power available.

Mathematically, gravity waves are solutions of the Navier–Stokes equation, the fundamental differential equation of fluid dynamics that forms the basis of all general circulation and numerical weather-prediction models. If gravity waves are omitted, time steps in numerical integration can be increased, and computer models run many times faster. A lot of ingenuity was applied to deriving and justifying hydrodynamic equations that precluded gravity-wave generation. Later observers recognized that ignoring such waves altogether worsens forecasts, especially during rapid weather transitions.

By the end of the 1960s, forecasters had reverted to using the full hydrodynamic equations, and all modern numerical weather models allow for gravity-wave generation. However, the idea of a simplified model that captures the useful properties of gravity waves without overloading computers still looks appealing.[6]

The beginning of the space age in the early 1950s also brought attention to gravity waves. Those were days of the Cold War, and both the Western and Eastern blocs were trying to build early warning systems for airborne attacks that relied on radio signals reflected by the ionosphere. Using rockets, scientists discovered that the atmosphere extends much higher than was previously thought. The first rockets were outfitted with aerosol ejectors that released columns of smoke in the upper atmosphere. The smoke was expected to expand by diffusion and to be displaced by blowing winds. Instead, the scientists saw strange and intricate distortions reminiscent of long-lived meteor contrails.

It soon became clear that the upper atmosphere is not smooth, but is continuously perturbed by irregular disturbances of various temporal and spatial scales, such as gravity waves, turbulence, and solar radiation and flares. Other inputs—energetic particles, the geomagnetic field, and electromagnetic forces—also cause distortions in the upper atmosphere. Interestingly, in the mid 20th century, aeronomists—scientists who study the upper atmosphere—were largely unaware of gravity waves, and meteorologists had written them off as insignificant noise. At the end of the 1950s, Colin Hines, an aeronomist, was among the first to link numerous observed atmospheric features to gravity waves propagating from below. He was struck by the coupling and later described it in everyday terms: "The ionospheric regions would be like a light-weight tail wagged by a very massive dog, and they must respond to almost any disturbance created below."[7]

The last two decades of the 20th century saw the proliferation of then revolutionary ground-based remote sensing techniques, including incoherent scatter radars based on reflection from density fluctuations, meteor radars exploiting ionized meteor trails, and pulsed lasers called lidars. Unlike rockets, those techniques do not disturb the atmosphere during observations and can almost continuously survey different altitudes up to about 100 km by analyzing reflected signals. The new measurements revealed many details about the structure and dynamics of the middle atmosphere and helped highlight the importance of gravity waves to atmospheric flows.

Perhaps the most often reported large-scale effect produced by smaller-scale waves is the creation of the coldest spot on

Earth, which is over the summer pole near the mesopause (see the article by Bodil Karlsson and Ted Shepherd, PHYSICS TODAY, June 2018, page 30). Gravity waves propagating from below grow in amplitude and break in the middle atmosphere, just like water waves break in a surf zone. The deposited wave momentum drives pole-to-pole air flow in the middle atmosphere. The flow is directed from the summer into the winter hemisphere, similar to the current along the shore produced by breaking surface waves. In addition to pole-to-pole flow, upward and downward currents also move over the summer and winter poles, respectively, and complete a circulation cell. The cooling produced by the adiabatic expansion of rising air over the summer pole exceeds the heating from an almost constantly illuminating Sun; thus that location becomes the coldest spot on Earth. On a summer evening at dusk, one can look to the north and see the high-altitude polar mesospheric clouds that occasionally form in the cold far reaches. They are often crisscrossed with traces left by gravity waves.[8]

Gravity-wave physics

Waves generated in the lower atmosphere by various sources have no preferred horizontal direction. The net horizontal momentum transported up and absorbed in the middle and upper atmospheres would thus be zero. So how can such an apparently chaotic field of waves create a well-defined pole-to-pole current on breaking? As waves travel vertically, they pass through circumpolar winds. If the horizontal phase velocity of a wave matches the persistent local wind, the distinction between the wave and flow disappears. The wave is then absorbed by the flow and can no longer propagate. Since the wind jets are aligned east–west in the summer hemisphere and opposite in the winter hemisphere, they selectively filter out some incident gravity waves.[9]

Many of the surviving waves that enter the middle atmosphere eventually break or cease their growth due to dissipation. Breakup is a violent process that begins with the instability of large-amplitude waves and is followed by a transition to turbulent motion. Therefore, the onset of breakup strongly depends on wave amplitude. Dissipation acts gradually, and over time the energy and momentum carried by waves are returned to the local flow. The behavior of the middle atmosphere is so influenced by breaking gravity waves that it cannot be explained without them.

The notion of the surf zone led some aeronomists to believe that all gravity waves break down in the middle atmosphere. That impression was indirectly supported by a surge of numerical global climate models whose upper bounds were exactly where the middle atmosphere ends, about 90 km above Earth's surface. In fact, the mesosphere is more like a reef surf zone where many waves break but larger-scale waves roll over and continue cruising toward the beach. However, the analogy with oceans ends there because the upper atmosphere beyond the mesosphere is viscous.

The molecular viscosity of a gas is proportional to the mean free path of the constituent molecules; it is therefore inversely proportional to the density and grows exponentially with height.

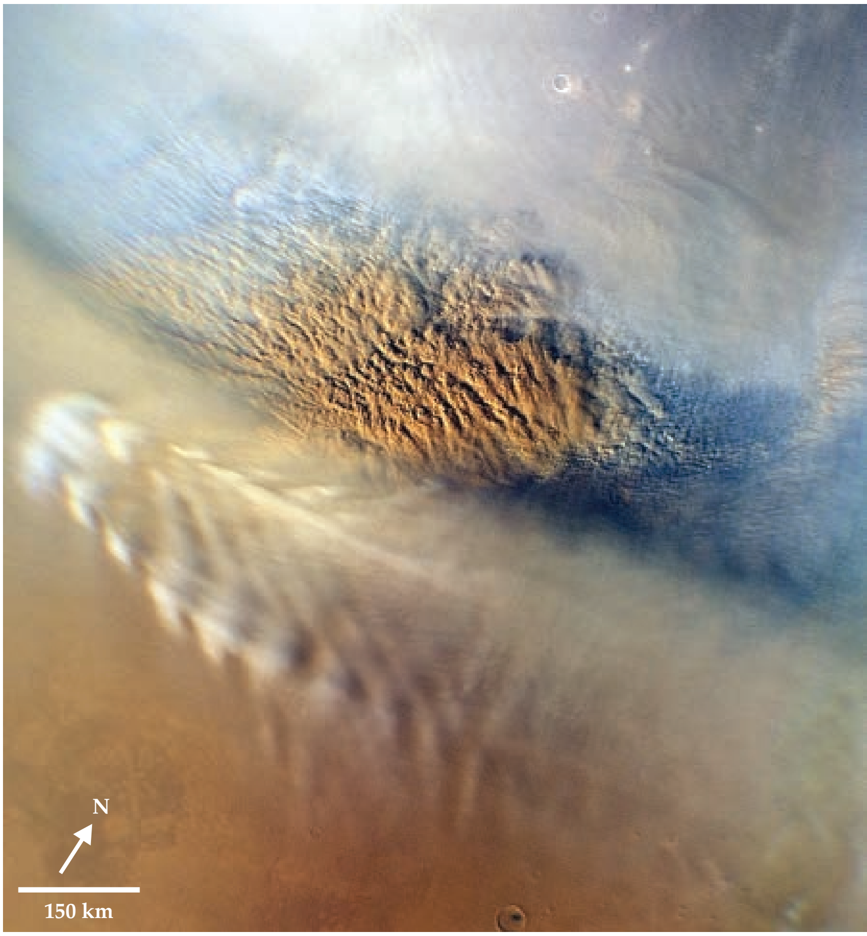


FIGURE 3. A MARTIAN DUST STORM was observed by the Mars Color Imager instrument on NASA's *Mars Reconnaissance Orbiter* on 7 November 2007 at around 3:00pm local time. This picture is centered on Utopia Planitia (53.6° N, 147.9° E); the seasonal polar north cap is seen at the top of the figure. Storms have variable durations; the dust storm pictured lasted for a day. Gravity-wave signatures are imprinted in the water ice clouds around Mie Crater in the bottom right of the image. The clouds form because of changes in atmospheric pressure and temperature from vertical air displacements caused by gravity waves that are excited by wind flow over a mountain. (Image courtesy of NASA/JPL-Caltech/MSSS. For more information, contact Malin Space Science Systems at www.msss.com.)

Competition between amplitude growth and exponentially increasing damping caused by molecular diffusion prevents gravity waves from breaking prematurely and helps some waves to propagate upward for several hundred kilometers and penetrate into what can already be considered space. The atmosphere as a continuous medium ends where collisions of molecules become so rare that their mean free paths greatly exceed spatial scales of motions. Thus the rarefied air at the upper edge of the atmosphere is already space for smaller-scale disturbances, but it can still sustain hydrodynamic waves with horizontal wavelengths of hundreds of kilometers.

The momentum transferred from breaking or dissipating waves to the local flow can accelerate or decelerate parcels of air in the upper atmosphere by hundreds of meters per second per day. That is a huge acceleration: Given that typical atmo-

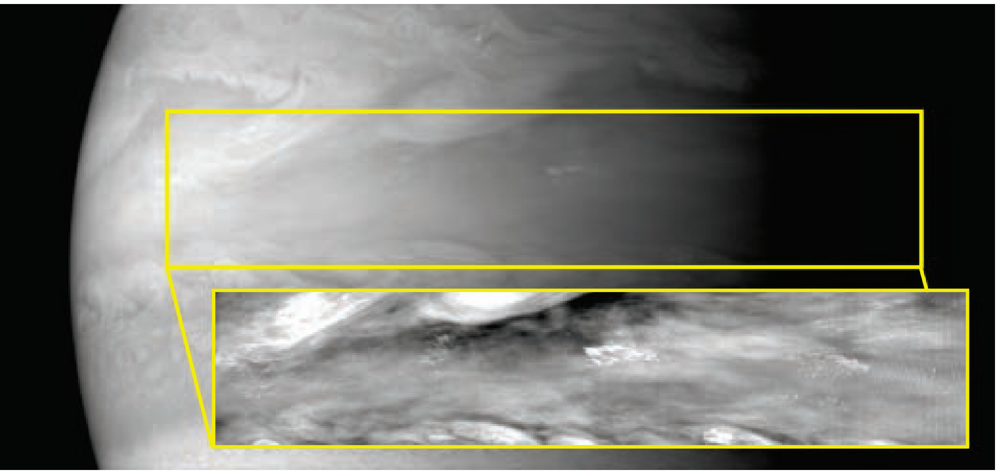


FIGURE 4. GRAVITY WAVES in Jupiter's equatorial atmosphere were detected in 2010 by the Multispectral Visible Imaging Camera aboard the *New Horizons* spacecraft. Jupiter is a gas giant and has no solid surface. Gravity waves are often generated by flow over mountains in terrestrial planets, but on Jupiter they are thought to be excited by convection deep below the visible clouds. Using the images from *New Horizons*, scientists found that the waves are moving much faster than the surrounding clouds. (Courtesy of NASA/Johns Hopkins University Applied Physics Laboratory/Southwest Research Institute.)

spheric wind velocities at those heights vary from several tens of meters per second to several hundreds, gravity-wave drag could kill them in about a day if no other forces were involved.

Gravity waves also affect the thermal state of the upper atmosphere. If a parcel of air is displaced up into a region with lower density, its temperature drops according to the first law of thermodynamics. Heating would then accompany its downward motion. In conservatively propagating waves, alternating phases of expansion and compression cancel each other out and produce no net input or loss of heat energy. On the other hand, when a wave attenuates with height, the expansions and compressions no longer cancel, and there is a net downward flux of sensible heat.[10] Thus, gravity waves and molecular thermal conduction work together to refrigerate and maintain the stability of the upper thermosphere.

The remarkable progress with satellite and ground-based observing systems, numerical models, and increased computational capabilities over the past decade has helped to uncover the importance of gravity waves in the middle and upper atmospheres. They constitute a major mechanism of vertical coupling by redistributing momentum and energy between atmospheric layers and driving many dynamical phenomena.[11]

Extraterrestrial atmospheres

Gravity waves are not unique to Earth; they can exist in all stably stratified fluids. Our solar system presents a great diversity of planets—including terrestrials, gas and ice giants, and dwarfs—and of atmospheres. Mars is the planet most visited by rovers and orbiters and the best known after Earth. Its thin atmosphere is reminiscent of Earth's mesosphere, but its rougher surface and stronger winds make its weather more volatile.

A thin, windy atmosphere and rough surface are also the ingredients for intensive gravity-wave generation. Remote sensing by orbiters confirmed that the gravity-wave activity on Mars is several times larger than on Earth, both in the lower and upper atmospheres. Figure 3 shows gravity-wave-induced variations in water-ice clouds as observed by NASA's *Mars Reconnaissance Orbiter* in 2007 during a local dust storm. Results from NASA's *Mars Atmosphere and Volatile Evolution* spacecraft suggest that the spectacular variability of the Martian thermosphere is produced by gravity waves traveling from sources near the planet's surface.[12]

Earth's closest planetary neighbor, Venus, has a scorching hot atmosphere due to its enhanced greenhouse effect. Thick clouds that cover the entire planet hide many details in the Venusian troposphere. Nevertheless, observations show clear wave signatures imprinted on the cloud top, and the Japanese *Akatsuki* orbiter recently detected very large scale gravity waves.[13]

Venus's atmosphere exhibits superrotation, a specific behavior in which the entire atmosphere rotates faster than the planet. Obviously, forces in the atmosphere must maintain that rotation, but their origin is unknown. Sophisticated numerical modeling provides ample indication that the mysterious force responsible for the superrotation is the torque delivered by gravity waves that originated in the atmosphere's hottest near-surface layers. As on Earth, they grow in amplitude with height and transfer their momentum to the mean flow on breaking at and above the cloud top.

Gravity waves have also been observed in the atmospheres of outer planets. In 1997 the *Galileo* probe, as it descended through the thermosphere and stratosphere of Jupiter, measured a temperature profile that revealed a gravity wavepacket propagating upward.[14] Recently, NASA's *Juno* spacecraft captured images of wave trains—more or less equally spaced crests and troughs—that exhibit gravity-wave behavior. (An image of Jupiter's atmosphere taken by *Juno* is shown on this issue's cover.) The upcoming *Jupiter Icy Moons Explorer* mission, which is scheduled for launch in 2022 by the European Space Agency, is expected to survey the Jovian system for more than three years. One of its prime objectives is to investigate the structure, dynamics, and composition of the Jovian atmosphere and thus provide more insight into gravity wave dynamics.

The *New Horizons* probe passed by Pluto in 2015 and discovered optically thin haze layers that extend to altitudes greater than 200 km. The most plausible mechanism for maintaining those clouds is related to gravity waves generated by diurnal sublimation and flow over topography.[15] On its way, the spacecraft also surveyed the Jovian system and captured an unprecedented view of gravity waves in Jupiter's equatorial atmosphere (see figure 4).

Living with gravity waves

Understanding the upper atmospheres of our own and other planets is necessary for learning how to navigate them. In the lower atmosphere, CAT challenges pilots and is thought to be produced by a superposition of many gravity waves. At higher altitudes, gravity waves can affect returning spacecraft, disturb satellite tracking, and skew GPS signals. Along with other larger-scale waves, they not only perturb atmospheric density at higher altitudes but also control the mean atmospheric state.

Space weather—the conditions surrounding Earth and other planets—is also remarkably coupled to the weather in the lower atmosphere by gravity waves. Space weather models are essential tools used to predict the Sun's effect on Earth. Geomagnetic storms are a fascinating manifestation of space weather caused by the interaction of solar energy and cosmic rays with Earth's magnetic field. During geomagnetic storms, the thermosphere–ionosphere system undergoes substantial changes, which can greatly affect propagation conditions of gravity waves. Thus, if we are to improve the predictive capabilities of space weather models, the influence of the whole spectrum of waves from below must be better understood and quantified.

On Mars and Venus, spacecraft perform aerobraking to modify their orbits by dipping into denser atmospheric layers of the thermosphere and mesosphere. Onboard accelerometers show that during every such maneuver, a spacecraft goes through a washboard-like profile of air density fluctuations. Understanding such gravity-wave-induced variability is important for planning and executing aerobraking operations, and for managing orbiters and ensuring the safety of their onboard instruments.

Continuing exploration and possible colonization of Mars will greatly increase the number of spacecraft that traverse the planet's atmosphere. As microsatellite architectures become more popular, Mars may be explored and monitored by swarms of individual satellites. Density variations associated with gravity waves can greatly affect the motions of small satellite swarms and larger spacecraft alike.

Although they may seem insignificant and obscure in the daily lives of humans, internal gravity waves are fundamental in atmospheric dynamics because they link vertical layers. Future observations of other planetary atmospheres and even stars will help to uncover how universal the phenomenon really is.

REFERENCES

1. D. C. Fritts, M. J. Alexander, *Rev. Geophys.* **41**, 1 (2003).
2. J. Lighthill, *Waves in Fluids*, Cambridge U. Press (1978).
3. E. E. Gossard, W. H. Hooke, *Waves in the Atmosphere: Atmospheric Infrasound and Gravity Waves, Their Generation and Propagation*, Elsevier (1975).
4. T. N. Venkatesh, J. Mathew, *Sādhanā* **38**, 707 (2013).
5. J. E. Fjeldstad, *Geofys. Publ.* **10**(6), 3 (1933).
6. J. Vanneste, *Annu. Rev. Fluid Mech.* **45**, 147 (2013).
7. C. O. Hines, *PAGEOPH* **130**, 151 (1989), p. 159.
8. M. Rapp et al., *J. Geophys. Res.* **107**, 4392 (2002).
9. D. G. Andrews, J. R. Holton, C. B. Leovy, *Middle Atmosphere Dynamics*, Academic Press (1987).
10. R. L. Walterscheid, *Geophys. Res. Lett.* **8**, 1235 (1981).
11. E. Yiğit, A. S. Medvedev, *Adv. Space Res.* **55**, 983 (2015).
12. E. Yiğit et al., *Geophys. Res. Lett.* **42**, 8993 (2015).
13. T. Fukuhara et al., *Nat. Geosci.* **10**, 85 (2017).
14. L. A. Young et al., *Science* **276**, 108 (1997).
15. G. R. Gladstone et al., *Science* **351**, 1280 (2016).
16. Y. Miyoshi et al., *J. Geophys. Res. Space Phys.* **123**, 2141 (2018). PT

Karl van Bibber is a professor of nuclear engineering and associate dean for research of the College of Engineering at the University of California, Berkeley. **Konrad Lehnert** is a professor of physics and JILA Fellow at the University of Colorado, Boulder, and the National Institute of Science and Technology. **Aaron Chou** is a senior scientist at the Fermi National Accelerator Laboratory in Batavia, Illinois.



Putting the squeeze ON AXIONS

Karl van Bibber,
Konrad Lehnert, and Aaron Chou

Microwave cavity experiments make a quantum leap in the search for the dark matter of the universe.

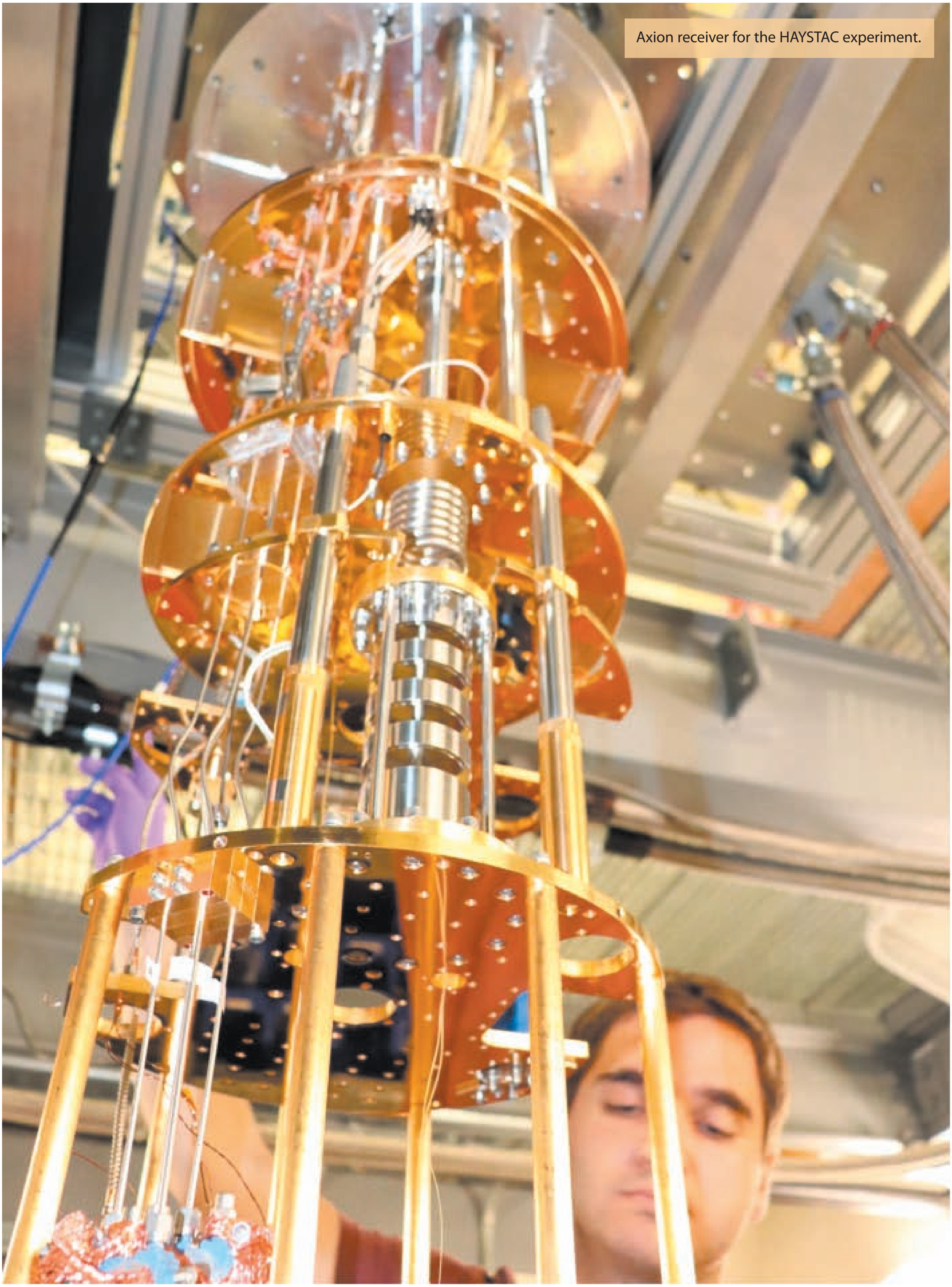
Sixty years ago Norman Ramsey and collaborators asserted that the neutron's electric dipole moment (EDM)—a measure of the separation of its positive and negative electric charge—was consistent with zero. More precisely, their experiment[1] bounded the neutron's EDM at less than 5×10^{-20} e·cm. Today, that limit is 3×10^{-26} e·cm, and experiments under development may push it lower by a factor of 100.

In the parlance of fundamental symmetries, the strong interaction is seemingly protected from *CP*-violating effects, where *CP* is the product operator of charge conjugation *C* and parity *P*. In the 1950s theorists had no compelling reason to expect *CP* violation—indeed, Tsung-Dao Lee and Chen Ning Yang did not believe that a nonzero neutron EDM would ever be found, though it was a worthy experimental question. However, with the advent of quantum chromodynamics (QCD) in the 1970s, a major problem loomed: The theory unavoidably includes a *CP*-violating angle θ associated with topological configurations of the QCD gluon field. For any generic value of θ, the neutron EDM should be a whopping 10^{-16} e·cm. The value implied by experiment is highly improbable: It's as if you spun a roulette wheel, and it came to rest at the winning number to within 1 part in 10 billion—just by dumb luck.

Or not? In 1977 Stanford University physicists Roberto Peccei and Helen Quinn conceived a minimal and appealing theory by which θ would be promoted to a dynamical variable. Just below some large energy scale, θ would assume a random value at each point in space. But in the low-energy limit of the theory, the nontrivial "washboard" potential of the QCD vacuum would drive θ to the *CP*-conserving minimum. That would be nice and tidy.

However, within a few months, Steven Weinberg and Frank Wilczek independently realized that the remnant sloshing of the θ field around that minimum implied the existence of an elementary particle called the axion—the smoking gun of Peccei and Quinn's theory—whose mass is possibly a trillion times lighter than an electron. Even such a light particle could, if sufficiently abundant, constitute the 27% of the mass–energy of the universe that is dark matter.

The story of the axion and its connection to dark matter is delightfully told with deep physical intuition in the form of a fable in which graduate students play snooker on Mars (see the article by Pierre Sikivie, PHYSICS TODAY, December 1996, page 22). In the current article, we focus on the ongoing experimental



Axion receiver for the HAYSTAC experiment.

hunt for the axion more than two decades later.

To date, the most sensitive searches rely on the fact that the axion couples to two photons. However, one may represent an external magnetic field as a sea of virtual photons, and as Pierre Sikivie realized in 1983, a massive axion can be converted into a single real photon in an external magnetic field with the same total energy. What's more, the axion–photon conversion can be resonantly enhanced in a high-Q cavity.[2] (See the article by van Bibber and Leslie Rosenberg, PHYSICS TODAY, August 2006, page 30.)

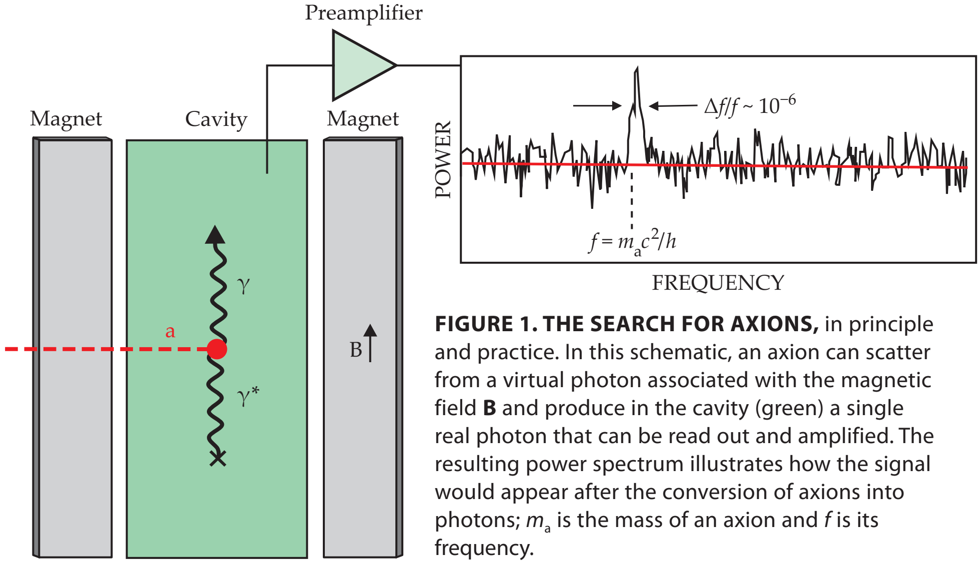


FIGURE 1. THE SEARCH FOR AXIONS, in principle and practice. In this schematic, an axion can scatter from a virtual photon associated with the magnetic field **B** and produce in the cavity (green) a single real photon that can be read out and amplified. The resulting power spectrum illustrates how the signal would appear after the conversion of axions into photons; m_a is the mass of an axion and f is its frequency.

Listening to the radio

Sikivie's proposed scheme was simplicity itself, and the experiments of today closely resemble larger and more technologically sophisticated incarnations of his first experiments 30 years ago. The new experiments boast a state-of-the-art low-noise amplifier that is coupled to a tunable microwave cavity inserted in the bore of a powerful superconducting solenoid magnet,[3] as shown in figure 1.

The cavity is tuned in small steps. At each frequency, the researchers pause for several minutes and listen for the signal—an excess of power over the noise floor—if the resonance condition is fulfilled, $h\nu = m_a c^2$. Here, m_a is the axion mass, ν the cavity frequency, and h Planck's constant, with 1 GHz corresponding to an axion mass of roughly 4 μeV. Think of the experiment as a revved-up version of your car's radio receiver.

The search strategy is dictated by the Dicke radiometer equation,

$$\frac{S}{N} = \frac{P_s}{k_B T_N} \cdot \sqrt{\frac{t}{\Delta\nu}},$$

where S/N is the signal-to-noise ratio, P_S is the signal power, k_B is Boltzmann's constant, and T_N is the system noise temperature. The factors under the square root are the integration time t at each step and the bandwidth of the axion signal $\Delta\nu$.

Although deceptively simple in concept, the experiment is one of the most daunting endeavors in physics today. Three factors complicate it. First, one must scan a range of axion masses over at least three decades. Because the search is narrowband, each decade must be covered by dint of many millions of tiny steps. Second, even for the most favorable axion–photon couplings predicted, and in the largest such experiment, the anticipated signal power is measured in units of yoctowatts, a trillionth of a trillionth of a watt. Third—and herein lies the real rub—unless one gets clever, a fundamental, irreducible noise floor set by quantum mechanics prevails for standard linear amplifiers.

Known as the standard quantum limit (SQL), the noise floor is expressed in terms of a temperature as $k_B T_{SQL} = h\nu$. More precisely, the system noise consists of a sum of two components, the familiar blackbody contribution (in parentheses below, where T is the physical temperature) and the noise accruing from the amplifier, T_A:

$$k_B T_N = h\nu \left(\frac{1}{e^{h\nu/k_B T} - 1} + \frac{1}{2} \right) + k_B T_A.$$

Half of the irreducible single quantum of noise $h\nu$ comes from the vacuum fluctuations of the cavity, even at zero temperature, and half comes from the linear amplifier itself. A convenient benchmark to keep in mind is that $k_B T_{SQL} \approx 50$ mK at 1 GHz. At high temperatures, $k_B T \gg h\nu$ and $T_N \approx T + T_A$.

The first microwave cavity experiments in the late 1980s at Brookhaven National Laboratory and the University of Florida used transistor-based amplifiers (heterojunction field-effect transistors that operated with system noise temperatures typically 5–20 K), some 200 times the standard quantum limit, T_{SQL}, over the 1–3 GHz frequency range. In the mid 1990s, the first large-scale experiment, the Axion Dark Matter Experiment (ADMX), began taking data. It used the best broadband amplifiers of its time; based on high-electron-mobility transistors, those amplifiers steadily improved the noise temperature to about 100 T_{SQL} at subgigahertz frequencies.

By virtue of its large volume cavity, ADMX had scanned a significant range in mass at one of two representative axion–photon couplings—corresponding to one variant of the KSVZ (named for Jihn Eui Kim, Mikhail Shifman, Arkady Vainshtein, and Valentin Zakharov) family of models regarded by the axion community as a useful experimental goalpost. But delving much deeper into the model space seemed out of reach. Furthermore, the scanning rate was unacceptably slow. Unless much better amplifiers were devised, the search for dark-matter axions seemed headed for an abrupt dead end.

Closing in on the quantum

The axion experiment unexpectedly gained a new lease on life with a chance conversation during a 1994 workshop at the University of California, Berkeley. Speaking to Leslie Rosenberg and one of us (van Bibber), then ADMX spokespersons, John Clarke ventured that he could make gigahertz amplifiers based on superconducting quantum interference devices (SQUIDs). At the time, amplifiers were limited to DC applications such as

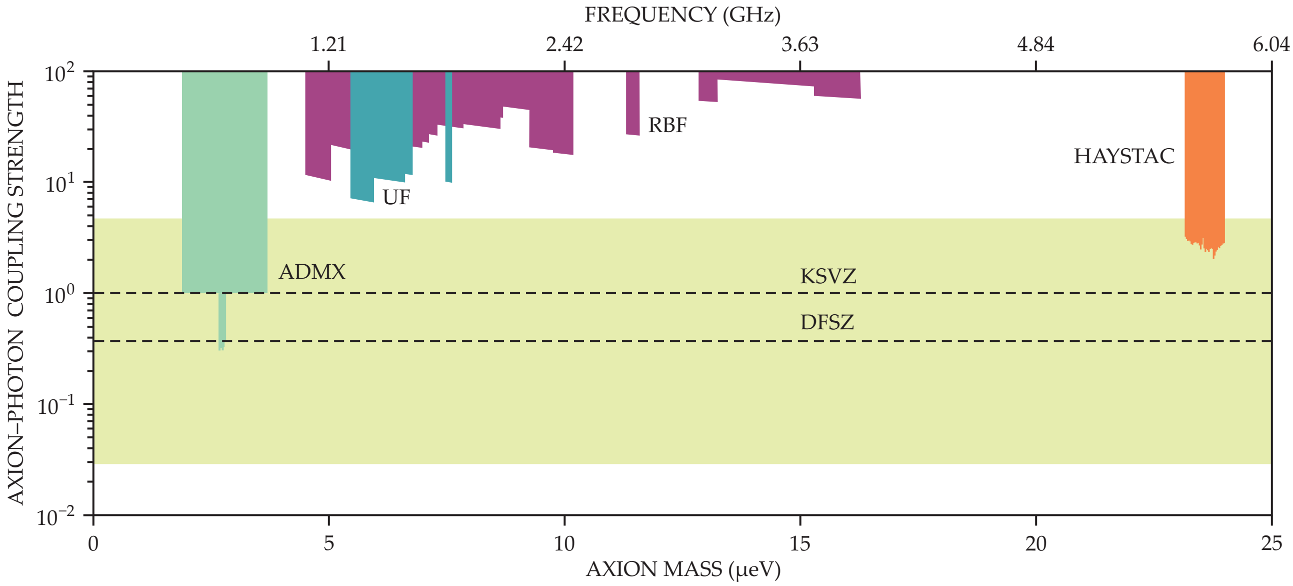


FIGURE 2. THE PARAMETER SPACE of axion mass and the axion–photon coupling strength excluded by microwave cavity axion experiments. The greenish yellow band delimits regions of the frequency spectrum already explored by theoretical models. The dashed lines are two specific realizations—from the KSVZ and DFSZ families of models—long used by the axion community as experimental benchmarks. UF and RBF denote first-generation experiments from the University of Florida and the Rochester-Brookhaven-Fermilab collaboration.[3] ADMX is the first large-scale US microwave cavity experiment, and HAYSTAC is the Yale-Berkeley-Colorado experiment. (Adapted from ref. 6.)

magnetometry. NSF took interest, and by 1998 Clarke and collaborators had invented the microstrip-coupled SQUID amplifier, or MSA. The achievement demonstrated quantum-limited performance on the bench.[4]

With a phased upgrade to MSAs and a dilution refrigerator approved by the Department of Energy, the experiment was finally able this past year to reach a more stringent goalpost in axion–photon coupling—one corresponding to a particular variant of the DFSZ (named for Michael Dine, Willy Fischler, Mark Srednicki, and Ariel Zhitnitsky) family of models. The system noise temperature[5] is now estimated to be only about 15 T_{SQL}. Just two years ago, a Yale-Berkeley-Colorado experiment called HAYSTAC published results in the 6 GHz (roughly 24 μeV) range. Using so-called Josephson parametric amplifiers (JPAs), the researchers demonstrated a system noise temperature a mere factor of 2 above the quantum limit and probed the model band (figure 2) with a cavity volume only 1% that of ADMX.[6]

One might think that further technology development would be unnecessary, and that only the size scale of the experiment and scan time need be extended for a definitive observational campaign. Chastened by fruitless decades hunting for the dark matter, however, the community of axion hunters should be prepared for a long march yet and have every technological advantage in their quiver.

Today's experiments are already pressing up against a fundamental limitation of quantum mechanics. But might there be a loophole that will enable us to continue improving the signal-to-noise ratio and scan speed of the search? In fact, quantum measurement is a rich and subtle topic, providing much room to maneuver. Both NSF and DOE have recently launched bold initiatives in quantum sensing and information in high-energy physics. They are supporting, in particular, two quantum-enhanced strategies in the search for axionic dark matter.

Putting the squeeze on

The first attempt to circumvent the standard quantum limit in the axion search is already under way at Yale University. In the summer of 2018, the HAYSTAC experiment began an upgrade, led by one of us (Lehnert), to a squeezed-vacuum state receiver, shown in figure 3. If successful, the project may be the first fundamental science experiment to use the squeezed states of the vacuum as data. (Incidentally, the Laser Interferometer Gravitational-Wave Observatory has prototyped an analogous system to improve how sensitively it detects gravity waves.)

Squeezing overcomes an apparent compromise in selecting the quality factor of the resonant cavity in the axion "radio." A higher-quality cavity boosts the amplitude of an axion-induced microwave tone because the tone accumulates for a longer time inside the cavity. But a higher-quality cavity also narrows the range of frequencies, or bandwidth κ, over which the radio achieves its best sensitivity. The narrowing would consequently require that the cavity be tuned in smaller steps.

It is never desirable to increase the rate κ_l at which the cavity absorbs energy. But in addition to that intrinsic loss rate, energy stored in the cavity also leaves at a rate κ_m through its measurement port. The optimum value of $\kappa_m = 2\kappa_l$ in a search for a signal of unknown frequency already sacrifices some sensitivity to a resonant signal for a larger overall bandwidth $\kappa = \kappa_l + \kappa_m$. Although the conflict between sensitivity and bandwidth may seem to be a mundane detail of microwave engineering, overcoming it requires techniques that are at the forefront of quantum measurement science.

Quantum squeezing maintains high sensitivity over a larger bandwidth by reducing one component of the cavity's quantum noise and noiselessly measuring that component, as sketched in figure 4a. The cavity mode that couples to the axion field is modeled as a single quantum harmonic oscillator, in which the cavity electric field is the oscillating quantity. After factoring out the rapidly oscillating contribution at the cavity's resonance frequency, the cavity field can be described by its slowly varying cosine X and sine Y components; the dimensionless variables X and Y are normalized by the scale of the cavity vacuum fluctuations.

Those "quadrature" variables, which follow the commutation

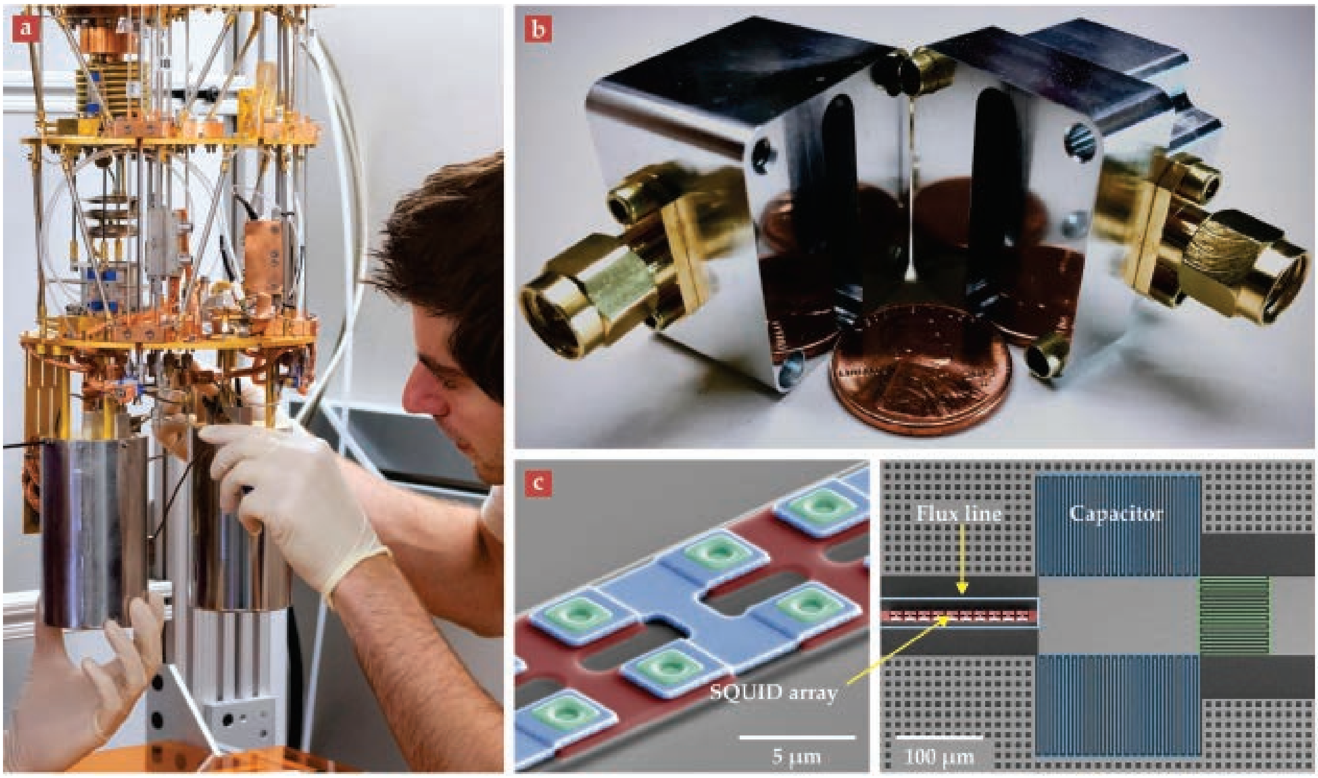


FIGURE 3. EXPERIMENTAL PROTOYPES. (a) Researchers assemble the two Josephson parametric amplifiers in this squeezed-state receiver for the HAYSTAC experiment. **(b)** The 7.1 GHz aluminum cavity for the squeezed-state receiver is split open. **(c)** This microphotograph shows a Josephson parametric amplifier composed of an array of superconducting quantum interference devices (SQUIDs). (Photographs courtesy of Dan Palken.)

relation $[X,Y] = i$, are constant under the evolution of the cavity Hamiltonian. The presence of an axion field can then be sensed through the fact that it would alter the state of the cavity in X,Y phase space. But unlike a classical description of a harmonic oscillator, a quantum oscillator cannot be localized in phase space with unbounded precision.

The axion cavity is continuously prepared in its ground state by virtue of the dilution refrigerator's cold environment. But zero-point (or vacuum) fluctuations $\Delta X^2 = \Delta Y^2 = ½$ ensure that the state is only localized to the minimum area consistent with the uncertainty principle $\Delta X{\cdot}\Delta Y = ½$. Simultaneous measurement of both X and Y must also add noise to preclude localizing the state beyond the Heisenberg uncertainty bound. Indeed, conventional amplifiers continuously measure both cavity quadratures when they boost the signal exiting the cavity. Repeatedly preparing the oscillator in its ground state and measuring its location in phase space results in values of X and Y that fluctuate[7] such that the apparent average energy of the oscillator is at least the standard quantum limit value of $h\nu$.

For the case of HAYSTAC, overcoming that SQL with squeezing is particularly natural because the JPAs already used in the experiment can noiselessly amplify one quadrature while squeezing the other. As outlined in figure 4, a first JPA prepares the cavity in a state with quantum fluctuations squeezed in X and equivalently amplified in Y. In a characteristic cavity storage time $t_s = 1/\kappa$, the cavity's state will be displaced by any axion field oscillating by an amount proportional to t_s near the cavity's resonance frequency. A second JPA then noiselessly amplifies and measures just the value of X. To the extent that the initial cavity state is arbitrarily squeezed in X and the subsequent measurement of X is noiseless, arbitrarily small axion displacements of the X component can be resolved.

A full analysis of the benefits of a squeezed-state receiver accounts for imperfections in the preparation of the squeezed state and in the single quadrature measurement and for the intrinsic loss of the axion cavity itself.[8] This last factor translates the notion of a measurement completely free of quantum noise into a practical increase in bandwidth and scan rate. In current experiments, the axion signal's coherence time is expected to be 10–100 times as long as a typical axion cavity storage time $1/\kappa$. If one knew the axion frequency, the signal-to-noise ratio would be maximized by bringing the cavity into resonance with the spectrally narrow axion signal and by choosing the cavity measuring rate to match its internal dissipation rate $\kappa_m = \kappa_l$.

At that "critical-coupling" condition, squeezing does not improve the signal-to-noise ratio of a signal exactly on resonance. The squeezed state decays back into an unsqueezed state during its storage in the lossy cavity. But by increasing κ_m above the critically coupled value, more of the squeezed state survives its now briefer time in the cavity and both the measurement noise and signal are reduced together, preserving their ratio. With arbitrarily large squeezing, the maximum signal-to-noise ratio can be maintained over a bandwidth much larger than the critically coupled value of $2\kappa_l$. In that way, squeezing benefits an axion search because it allows the cavity to be tuned in larger frequency steps; a particular frequency range can thus be scanned in a shorter time.

The first implementation of a squeezed-vacuum state receiver on HAYSTAC will use that strategy. In a test at JILA, the Lehnert group has demonstrated squeezing the noise variance by a factor of −4.5 dB (data shown in figure 4b). And in a mock axion search experiment, the group sped up the scan rate by a factor of 2.1, exactly in accord with predictions for the system.[9] The twin of that system is currently being commissioned in HAYSTAC with a target of 2–3 increase in scan rate. A new squeezed-state receiver design is being explored at JILA that may speed up the scan rate by a factor of 10.

Zero to one

Instead of measuring the cavity wave in the X,Y quadrature basis, one can instead measure its amplitude and phase in polar coordinates, which are themselves conjugate observables satisfying the Heisenberg uncertainty principle. The axion signal would appear as a nonvanishing cavity electric field amplitude, as evidenced by the occasional appearance of a single photon in the otherwise unoccupied cavity. That signal would be a zero-to-one transition, much like a spurious qubit error in a modern quantum computer.

A longer-term and more ambitious goal to beat the standard quantum limit will be the implementation of a photon-counting detector that measures the amplitude but not the phase of the cavity wave. Just as in the case of quadrature squeezing described above, only one of these conjugate observables is measured, and so there is no fundamental limit on the measurement noise due to quantum back action. As Yale's Steve Lamoreaux and colleagues have pointed out, the measurement errors that arise from a photon-counting detector's dark-count background rate—that is, its spurious noise hits—could therefore be made arbitrarily low. The result would be pristine resolution of tiny field amplitudes for ultrasensitive axion searches.[10]

A joint Fermilab and University of Chicago team led by one of us (Chou), Daniel Bowring, and David Schuster is conducting precisely such R&D to lay the groundwork for an ambitious axion search at masses above 40 μeV (above 10 GHz), utilizing novel microwave photon-counting detectors based on artificial atoms—superconducting qubits developed for quantum computing.

The new microwave photon-counting detectors employ quantum nondemolition (QND) measurements, which repeatedly probe the electric field of the photon stored in the cavity without actually absorbing and thus destroying the photon in the process. These measurements rely on the atom's electric polarizability, a quantity that describes the potential energy associated with the off-resonance dipole scattering of a photon by the atom—the same scattering process that makes the sky blue. The interaction energy associated with the repeated scattering by even a single photon confined in a cavity creates an observable shift in the atomic energy levels.

The atom may be considered a ball-and-spring oscillator with the atomic electron represented by the ball and the nonlinear spring represented by the anharmonic $1/R$ Coulomb potential. The rms electric field of the background photon stretches the spring and exercises its nonlinearity, thus causing the resonant frequency of the atomic oscillator to change. For sufficiently strong coupling, the electric field of even a single photon can be resolved by that nondestructive amplitude-to-frequency transducer.

The nonlinear response is responsible for the Lamb shift, which is due to interactions of the atom with zero-point photon fluctuations of the quantum vacuum. In the case of interactions with photon modes of the finite, nonzero occupation number, the corresponding effect is known as the AC Stark shift. Just as in the case of the Lamb shift, the atom absorbs no net photons; the atom thus acts as a nondestructive photon sensor.

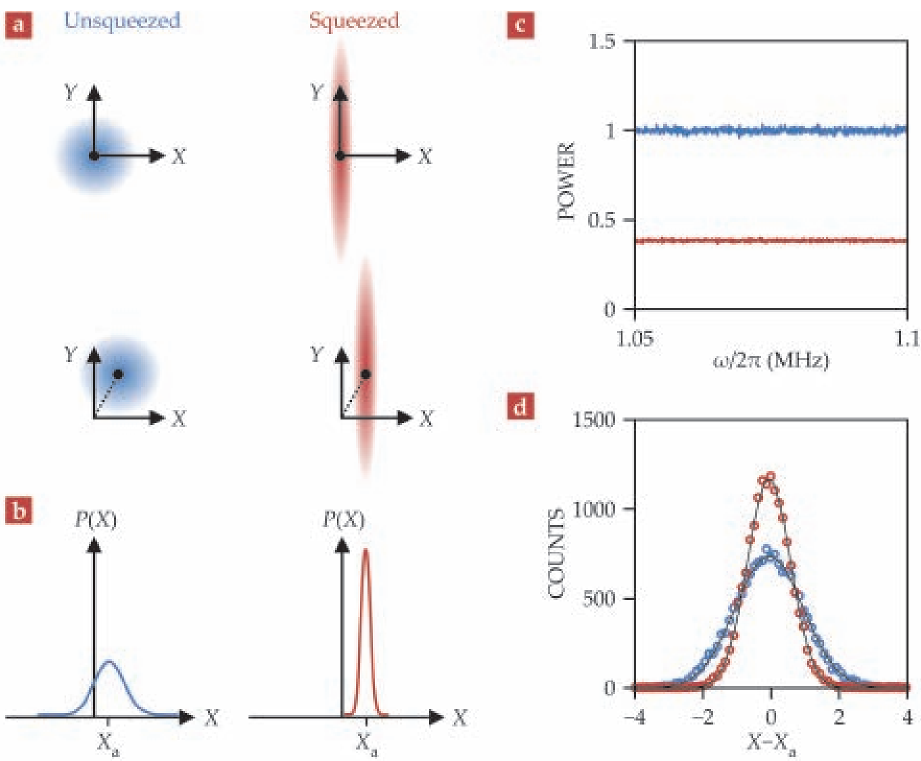


FIGURE 4. AN AXION SEARCH MEASUREMENT. In this illustration **(a)** of phase-space variables X and Y, the variables' noncommutation imply that the phase space cannot be localized to an area smaller than a Heisenberg uncertainty region. The cavity state is initially prepared in either its ground state (blue) or a squeezed state (red). But once an axion has entered the cavity, the state is displaced (the dotted line) by the axion field. **(b)** A noiseless measurement of the X component yields a probability density $P(X)$. Because noise has been squeezed from the X to the Y variable, the displacement in X by the actual axion signal is more easily detected. **(c)** Noise power from a squeezed-state receiver prototype is plotted versus frequency ω. **(d)** A histogram of measured values of X with (red) and without (blue) squeezing match the theoretical plot of panel b, in units of the vacuum noise.

By performing repeated spectroscopic measurements of the atomic transition frequencies, one can determine the exact photon occupation number of the cavity state and thus ascertain the presence or absence of the putative signal electric field. More specifically, in an axion search experiment, the cavity is cooled to its vacuum state and one measures how often the QND process observes a frequency shift corresponding to an $N = 1$ cavity state relative to the rate of observing the $N = 0$ vacuum state. The axion signal would then be a significant excess of photon counts above statistical fluctuations in the background counts (see figure 5). The methodology is equivalent to searching for excess narrowband power in signals read out with a linear amplifier. But the background dark count probability can now be reduced to a small fraction of the single photon per readout required by the standard quantum limit.

Although the measurement process conserves the cavity photon number, the atomic spectroscopy requires the atom to absorb probe photons at the shifted transition frequency. The increased atom–cavity interaction energy due to the extra absorbed probe photon causes a reciprocal shift of the cavity mode frequency. It thus also creates phase noise in the cavity photon state as measurement back action. In the limit of perfect QND measurement, the cavity state is projected into a state of definite photon number and maximally randomized and indefinite phase.

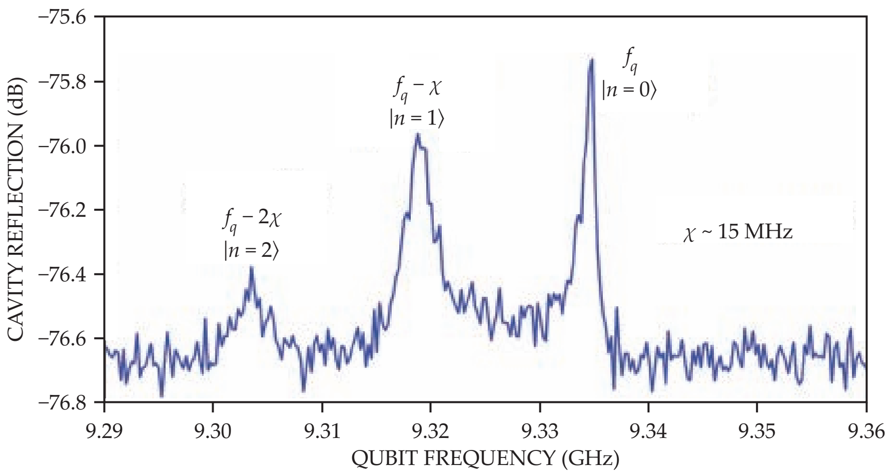


FIGURE 5. THE QUBIT RESPONSE SPECTRUM, measured when the cavity is driven with mean photon occupation number $\langle N \rangle \approx 1$. Due to the potential energy of interaction with a photon's electric field, the qubit's frequency f_q shifts by a quantized amount $\chi = 15$ MHz for each photon present in the cavity. The resulting spectrum exhibits a Poisson distribution that describes statistical fluctuations in the cavity occupation number. For the much smaller signals expected from axions, one would obtain a Poisson distribution with $\langle N \rangle \ll 1$, and the signature would be a single photon accompanied by a single unit of frequency shift. That signature would appear as the occasional population of the $N = 1$ peak of the spectrum. (Data courtesy of Akash Dixit.)

Any anharmonic oscillator with a nonlinear restoring force can act as the amplitude-to-frequency transducer necessary to implement a QND measurement. Physics Nobel laureate Serge Haroche and collaborators originally demonstrated the QND technique using beams of Rydberg atoms[11]—huge atoms of very large principal quantum number—but newer implementations exploit artificial atoms made of superconducting qubits.[12] In those qubits, the nonlinear kinetic inductance of a Josephson junction combines with the junction capacitance to form an anharmonic *LC* oscillator.

In practice, additional capacitance is added to reduce the susceptibility to charge noise, and that capacitance can be used to tune the resonant frequency of the circuit to any desired value. Also, unlike in real atoms, the strength of the qubit's electric dipole coupling to the background photon's electric field can be increased simply by attaching appropriately sized antennae to opposite sides of the capacitive junction. Again, the cavity photon's field can be thought of as an oscillating force that nonlinearly stretches the electromagnetic spring and changes its resonant frequency. The qubit-based QND sensor can be easily mounted inside the cavity on a dielectric substrate, as shown in figure 6.

Because the low temperatures achieved by dilution refrigerators suppress the cavity's blackbody photon population, the dark count rate in state-of-the-art superconducting qubit sensors is primarily determined by spurious readout errors. It's still a mystery why the qubits are found in their excited state far more often than are expected at the operating temperature. That situation produces false positives in the absorption spectroscopy used to probe the qubit frequency shifts. Typically, the probability of a spontaneously excited state is around 1%, which corresponds to an average of 0.01 dark counts per readout.

Although the resulting dark rate is less than the effective 1 dark count per readout for the standard quantum limit, the rate could still be improved. Efforts are under way to implement simultaneous or sequential QND readout of the cavity occupation number using gangs of independent qubit sensors and requiring all or a large fraction of them to report the same answer. Confirmation of a detection by many qubits should significantly reduce the resulting dark count rate below that from individual qubit readout errors.

More than a decade ago, Seishi Matsuki and collaborators on the CARRACK axion experiment in Kyoto, Japan, developed a single-photon detector using Rydberg atoms to resonantly absorb single microwave photons. It convincingly demonstrated a thermal-photon dark rate a factor of two below the standard quantum limit at 2.5 GHz;[13] an axion search was conducted over 10% bandwidth in mass with roughly DFSZ sensitivity, but it was never published. The QND techniques now being developed can achieve much greater fidelity in signal readout than that absorptive technique by making multiple redundant measurements, and the background photon rates can be dramatically reduced with subkelvin operation. A background-free axion experiment is potentially within reach.

Controlling the quantum

More than half a century ago, Chester Gould's comic strip detective Dick Tracy famously predicted, "The nation that controls magnetism will control the universe." That prediction was off

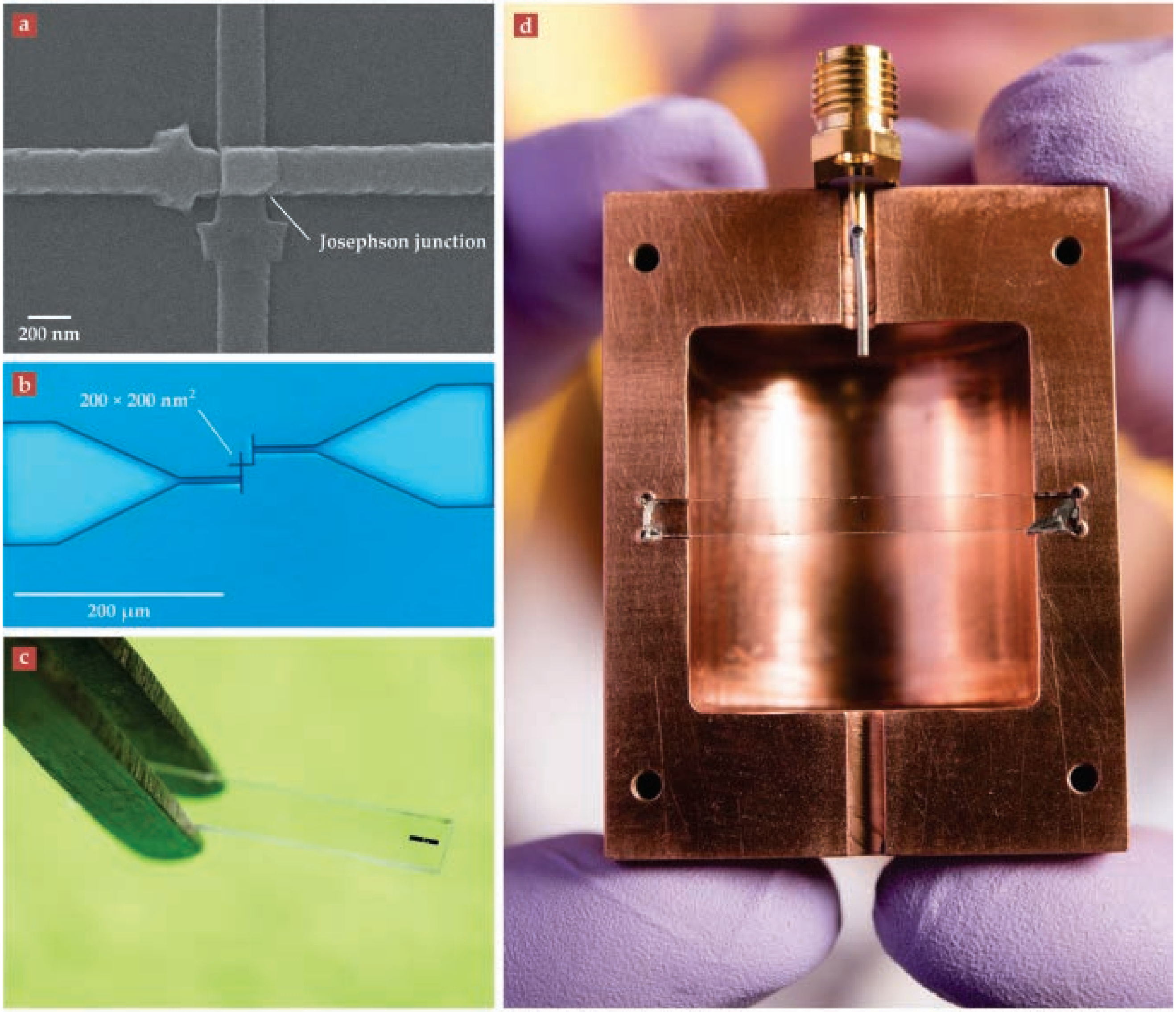


FIGURE 6. A SUPERCONDUCTING "ARTIFICIAL ATOM" QUBIT is an anharmonic *LC* oscillator **(a)** that uses the nonlinear inductance of a Josephson junction. **(b)** Larger superconducting structures may be attached to the junction to build up millimeter-size antennae **(c)**, which enable stronger coupling to the electric field of centimeter-wave cavity photons. **(d)** The qubit is mounted inside a cavity with a dielectric substrate. The vertically oriented electric field of a single-cavity photon "stretches" the qubit oscillator and exercises its nonlinearity. The quantum nondemolition photon detection protocol can be phrased as a yes–no question: Has the qubit's resonant frequency shifted in response to the appearance of a cavity photon or not? (Photographs courtesy of Akash Dixit and Reidar Hahn/Fermilab.)

the mark, perhaps, although Gould was spot on with his two-way video wristwatch. Controlling the quantum could be a different story, and we predict that the quantum will ultimately bring the dark matter of the universe into view.

Dramatically improved receivers based on quantum sensing are no panacea for the axion experiment. A parallel challenge not discussed in this article is the development of innovative microwave cavities satisfying multiple constraints of the axion experiment. Technologies that are being pursued include photonic bandgap resonators and the use of metamaterials and thin-film superconductors. But that's a story for another time.

The authors gratefully acknowledge support from NSF, the US Department of Energy, and the Heising-Simons Foundation.

REFERENCES

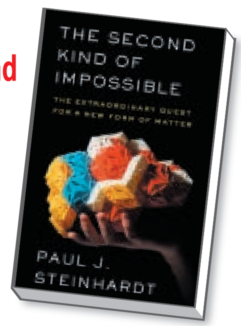
1. J. H. Smith, E. M. Purcell, N. F. Ramsey, *Phys. Rev.* **108**, 120 (1957).
2. P. Sikivie, *Phys. Rev. Lett.* **51**, 1415 (1983); erratum *Phys. Rev. Lett.* **52**, 695 (1984).
3. P. W. Graham et al., *Annu. Rev. Nucl. Part. Sci.* **65**, 485 (2015).
4. M. Mück et al., *Appl. Phys. Lett.* **72**, 2885 (1998).
5. N. Du et al. (ADMX collaboration), *Phys. Rev. Lett.* **120**, 151301 (2018).
6. B. Brubaker et al., *Phys. Rev. Lett.* **118**, 061302 (2017).
7. C. M. Caves, *Phys. Rev. D* **26**, 1817 (1982).
8. H. Zheng et al., https://arxiv.org/abs/1607.02529v2.
9. M. Malnou et al., *Phys. Rev. X* **9**, 021023 (2019).
10. S. K. Lamoreaux et al., *Phys. Rev. D* **88**, 035020 (2013).
11. S. Gleyzes et al., *Nature* **446**, 297 (2007).
12. B. R. Johnson et al., *Nat. Phys.* **6**, 663 (2010).
13. M. Tada et al., *Phys. Lett. A* **349**, 488 (2006).

PT

© CHRIS BOLIN

Aperiodic Penrose tiling at Oxford University.

The hunt for natural quasicrystals



The Second Kind of Impossible

The Extraordinary Quest for a New Form of Matter

Paul J. Steinhardt
Simon & Schuster, 2019. $27.00

Scientific discovery is the successful extension of human knowledge into the unknown. Particularly satisfying are those discoveries that appear impossible at first glance but upon closer inspection reveal a long-overlooked loophole in our most basic assumptions. *The Second Kind of Impossible: The Extraordinary Quest for a New Form of Matter* by Paul Steinhardt is the story of several such gaps in the fields of crystallography and mineralogy. It is a thrilling mix of scientific memoir and true detective story. Most importantly, it is a tale of the excitement that drove the author to extraordinary insights far outside his original area of expertise.

Quasicrystals, the main subject of the book, are complex ordered phases of matter first found in metallic alloys. Their study has a rich history filled with unique protagonists and unexpected developments. Thus they are a perfect topic for a scientifically themed book aimed at a general audience.

The first part of *The Second Kind of Impossible* introduces the foundations of quasicrystal research. It covers the origins of crystallography in the 18th century through the end of the 20th century, when quasicrystals became established as a real and fascinating form of matter. In the 1970s mathematician and physicist Roger Penrose introduced Penrose tilings, intricate and aesthetically beautiful patterns with long-range order but without periodicity. The tilings were thought to be a mathematical gimmick until two independent advancements brought them to the attention of a wider community. At Princeton University in the early 1980s, Steinhardt and his graduate student Dov Levine envisioned an icosahedral generalization of Penrose tilings. They dubbed that hypothetical new form of matter quasicrystals. Unbeknownst to Steinhardt, materials scientist Daniel Shechtman, using electron microscopy, had already identified an aluminum-manganese alloy with icosahedral diffraction symmetry.

Steinhardt describes both the mathematics that went into his theory and his construction of geometric models from styrofoam balls, pipe cleaners, and paper. The astonishingly fruitful interplay between theory and model building soon explained all aspects of Shechtman's discovery and subsequent developments. After that success, Steinhardt needed a new challenge to continue his research. As of 1999 all known quasicrystals were synthetic in origin, so Steinhardt began wondering if perfect quasicrystals might develop without human intervention.

The book next dives into Steinhardt's 20-year hunt to answer that question. He reasoned that if a natural quasicrystal existed, it could be found in some known but incompletely characterized mineral sample. But as a theoretical physicist and cosmologist, he had no formal training in mineralogy. He contacted experts in materials characterization around the world to ask for their assistance and collaboration.

The reader follows Steinhardt as he recounts scouting crystal diffraction databases and exploring the intricacies of microscopic crystalline grains. Convincing colleagues of his unexpected findings was not always easy. Against the odds and after many drawbacks, including dealing with smugglers and deciphering clues in secret diaries, Steinhardt and his closest collaborator, Italian geologist Luca Bindi, traced the origins of a natural quasicrystal candidate to one of the most remote places on Earth: the Koryak Mountains in Siberia.

In the final part of the book, Steinhardt and a team of 12 scientists, prospectors, and support personnel make an adventurous field expedition to the Koryak Mountains. There they hope to find needles in a haystack—tiny rocks only millimeters in size, surrounded by thick clay in the banks of a pristine river bed. The samples they eventually uncover are of extraterrestrial origin and date back to the beginning of the solar system. I refrain from recounting the many astonishing turns of events the team encountered. But suffice to say that a fiction writer could hardly have thought of better plot twists. This section was the highlight of *The Second Kind of Impossible*. The hunt captured and held my attention; I could not put the book down.

One of the book's strengths is its accessibility. Despite the specialized topic, Steinhardt conveys with engaging passion his motivation and how it changed over the course of his quest. He describes his struggles with unreliable sources, competitors, and skeptics. His approach—assembling a "red team" of critics and a "blue team" of advocates that engaged in friendly competition until the scientific truth was revealed—is a formidable demonstration of how to avoid confirmation bias. Finally, and this I find a particularly important moral, Steinhardt uses his personal perspective to demonstrate that scientific discovery is often not a solitary effort. Instead, true progress comes from openness to the world and the acceptance of potential failure. The events described in the book are an extraordinary display of tenacity and serendipity, and the writing is captivating, entertaining, and full of fascinating scientific content. I strongly recommend *The Second Kind of Impossible* to experts and lay audiences alike.

Michael Engel
Friedrich-Alexander University Erlangen-Nuremberg, Germany

COLLECTION OF HISTORICAL SCIENTIFIC INSTRUMENTS, HARVARD UNIVERSITY

A pocket watch made for the Japanese market circa 1715.

Telling time in Tokugawa Japan

Making Time
Astronomical Time Measurement in Tokugawa Japan
Yulia Frumer
U. Chicago Press, 2018. $45.00

Yulia Frumer's *Making Time: Astronomical Time Measurement in Tokugawa Japan* will fascinate readers with its study of the evolution of different systems of time measurement in Japan. The book begins with the arrival of the first Western mechanical clocks during the Tokugawa (or Edo) period from 1603 to 1868, when the ruling shogunate restricted contact with foreigners and strengthened the country's traditional temporal system. The author then traces timekeeping technologies and systems through to the calendar reform of 1873, when the Japanese people wholeheartedly embraced Western methods of timekeeping.

In 17th-century Japan, Western clocks seemed nonsensical to consumers because they measured time in 24 equal hours and were dissociated from natural events like dawn and dusk. By contrast, the Japanese people divided their day into 12 unequal hours, following an ancient Chinese system introduced in Japan in the 7th century. Daylight and darkness were each split into six equal parts, no matter what the season. The daylight hours grew longer in the summer and shorter in the winter, with the reverse occurring for the nighttime hours. The hours of the day and night were equal in length only on the equinoxes. Each hour was named for 1 of 12 animal signs and was also given a number that could be used to announce the hour by strikes of a bell. Midday, for example, was Horse (9). A vestige of this system is still found in the modern Japanese expression that refers to morning as "before Horse" and afternoon as "after Horse."

Although sundials could find the time directly from the Sun overhead, water and fire clocks were commonly used to keep time. Incense clocks, a type of fire clock widely used in the Tokugawa period, had moveable hour markers placed in the sand alongside a trail of burning incense and required a standardized system for shifting the hour markers as the year progressed. The lengths of day and night hours were adjusted 24 times a year according to seasonal weather changes such as with the "beginning of spring," "rain water," "major heat," "cold dew," "frost descending," and "major snow."

Japanese users noticed that those seasonal weather markers were not in tune with Japan but with northern China, where the calendar had originated. Moreover, scholars saw the need to regulate the lunar calendar against the solar year so that the seasons would not drift. To that end, the central government established an astronomical bureau to manage and reform the calendar as needed.

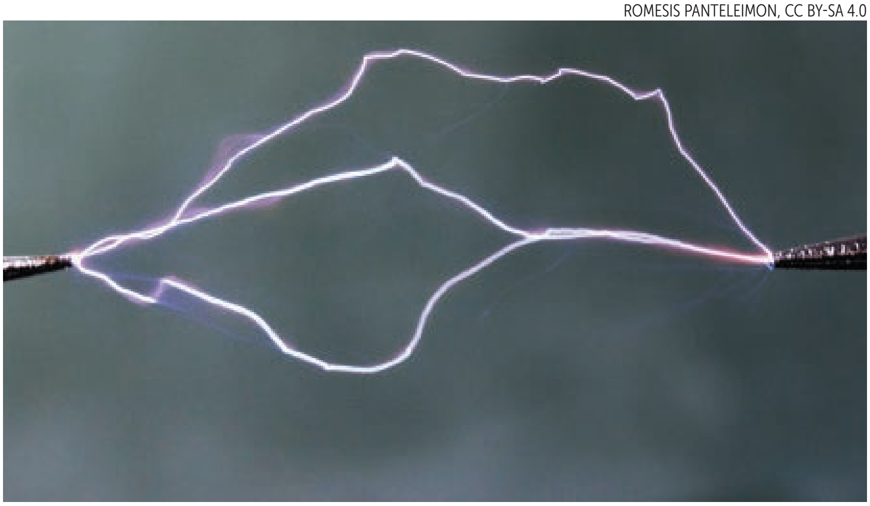
Frumer notes that by the late 17th century, Japan had in place a standardized calendar and a single time zone for the entire country. She also shows how Japanese clockmakers reengineered Western mechanical clocks to keep variable hours and shape them to suit Japanese life. Weights on the ends of the foliot in early mechanical clocks were shifted as often as twice daily to speed up or slow down the passage of day or night hours. Later mechanical clocks had moveable digits, index arms with adjustable lengths, and faces marked with hour lines like sundials.

Frumer argues that Tokugawa astronomers became comfortable with Western clocks and the system of 24 equal hours during the 18th century, as they measured the motion of heavenly bodies along celestial arcs and timed star transits. In the context of longitude determination, the author says, astronomers came to see time as a mathematical "belt" wrapping around the globe and the universe behaving more like an equal-hour timepiece than a variable clock of the early Tokugawa period. By the Meiji period, when Japanese society opened up to international commerce, Western timekeeping practices had become associated with positive values such as convenience, social progress, science, and enlightenment. The equal-hour system had another advantage for the Meiji government: It shifted people away from divination practices and superstitions linked to the unequal hours and their animal associations. Thus, in 1873 the Japanese government mandated the use of the 24-hour clock.

Making Time is the most comprehensive treatment of Japanese timekeeping to date, but it is not a specialized book for horologists interested in detailed information about clock mechanisms and makers. Frumer's text is addressed to historians of science, technology, and Japanese culture. She deftly shows that technology is not just about practical needs; it is shaped by a society's values and activities. Frumer's book also prompts questions about technology transfer by showing how clocks from Europe that landed in Japan did not always have the same interpretation or use in their new environment. If the book has a shortcoming, it is that the narrow focus on timekeeping left me wishing for a broad history of Japanese methods of time finding and material culture. (I say this, admittedly, as the curator of a large collection of Japanese sundials.)

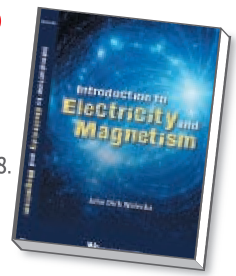
Time pluralism was also common in Europe during that time period, although Frumer doesn't make the point. Not only were the Gregorian and Julian calendars in simultaneous use, but solar time was read according to multiple systems of equal hours, which varied by geographical region and type of work, and the Catholic church also employed some timekeeping methods that used unequal hours. European pocket sundials enabled users to track the time in different systems, find the lengths of day and night in different seasons, and compensate for diverse latitudes. Those sundials shared similar functions with the clocks and astronomical activities discussed in Frumer's book, and their use was shaped by social needs in similar ways. That is not a criticism of the book; it is a note that Frumer's analysis has reach far beyond Japan.

Sara J. Schechner
Harvard University
Cambridge, Massachusetts



Electromagnetism textbook bridges the gap between basic and advanced

Introduction to Electricity and Magnetism
John Dirk Walecka
World Scientific, 2018. $48.00 (paper)



Several widely adopted physics textbooks in electricity and magnetism are available for undergraduates; they include *Introduction to Electrodynamics* by David Griffiths (4th edition, 2017) and *Electricity and Magnetism* by Edward Purcell and David Morin (3rd edition, 2013). For instructors and students hoping to supplement those books with a mathematically concise treatment, however, *Introduction to Electricity and Magnetism* by John Dirk Walecka would be a good companion text. A succinct treatment of the major topics in the field, it provides a bridge between basic and advanced textbooks.

Walecka is a well-known nuclear theorist and author of several undergraduate and graduate textbooks in nuclear physics, modern physics, and general relativity. *Introduction to Electricity and Magnetism* is divided into three sections and seamlessly transitions from electricity to magnetism and then to Maxwell's equations. Each section ends with a summary chapter of major concepts, a feature that readers will likely find useful. The assumption that the reader has already studied introductory classical mechanics and vector calculus enables Walecka to tackle more sophisticated mathematics than most introductions to the subject.

The first section covers the fundamentals of electricity, including Coulomb's law and electric fields, Gauss's law, electrostatic potential and energy, capacitors, and Ohm's law and circuits. Before moving on to topics in magnetism, Walecka gives a brief review of vectors, differential operators, and the Stokes theorem. Any reader wanting to review those topics will find the summaries a welcome feature.

The second section provides a good treatment of magnetism that covers magnetic force and fields, Ampere's law, induction, magnetic materials, and time-dependent circuits. Walecka constructs an analogy between electric fields and magnetic fields that is a nice segue between the first and second sections. He also incorporates illustrative examples throughout the book. For instance, when he introduces Ampere's law, he provides three excellent applications of it: magnetic fields due to an infinite sheet of currents, two opposite sheets, and an infinite solenoid. All three help readers grapple with what is often a challenging concept.

The third and final section of the book engages with Maxwell's equations. Walecka does a fine job of explaining how the incompleteness of Ampere's law prompts the need for Maxwell's equations. Additionally, the equations are expressed in both integral and differential forms, and Walecka shows the reader how to derive the electromagnetic wave equation from Maxwell's equations in a clear and intuitive way. The book concludes with a chapter on special relativity that will inspire readers to learn more about that topic.

The book's brevity helps readers see how seemingly different topics in electricity and magnetism connect with one another. The notations and illustrations are visually appealing and uncluttered. Moreover, the softcover book is affordable compared with most textbooks, a feature that students will surely appreciate.

Although the book's conciseness is a strength, it is also in some ways a weakness. Some readers may find *Introduction to Electricity and Magnetism* a bit formal and abstract. Diligent readers who are looking for conceptual explanations may need to refer to one of the introductory texts mentioned earlier. There are homework problems at the back of Walecka's book, but they are often not of a sufficient number for students to learn the material comprehensively. Additionally, readers may find themselves lost in the numerous derivations because the book does not refer to many real-life applications. Students new to a field often find real-life examples particularly illuminating, and I found myself wishing they had been included in the text.

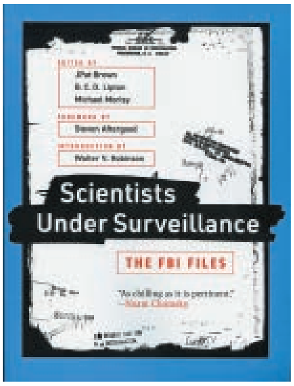
Overall, *Introduction to Electricity and Magnetism* is an excellent, concise introduction to the topic. It presents mathematical treatments of abstract concepts in a clear and straightforward way. I think it will be most effective as a companion to other excellent introductory texts, but readers who want to review the material will find Walecka's treatment of electricity and magnetism refreshing.

Savan Kharel
Williams College
Williamstown, Massachusetts

NEW BOOKS & MEDIA

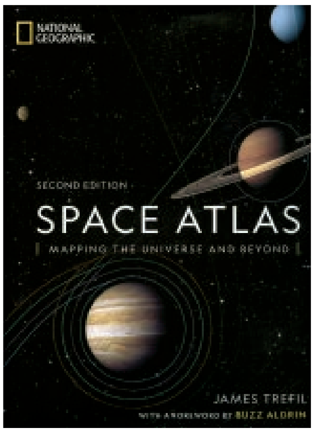
Scientists Under Surveillance

The FBI Files



JPat Brown, B. C. D. Lipton, and Michael Morisy, eds.
MIT Press, 2019. $24.95 (paper)

Throughout the Cold War, FBI director J. Edgar Hoover kept tabs on some of the most renowned scientists in the country. Nonprofit news site MuckRock filed thousands of Freedom of Information Act requests to obtain the files of figures such as Hans Bethe, Richard Feynman, Carl Sagan, and Vera Rubin, and MIT Press has now made them available in paperback. Many of the documents in *Scientists Under Surveillance* are heavily redacted, but the collection still makes for fascinating reading. The level of detail is often unnerving. For example, the FBI tracked every train, airline, and hotel reservation for the visiting Hungarian mathematician Paul Erdős on the following grounds: "While not identified as a communist, Erdos has maintained his ties with communist Hungary, allegedly to protect the pension being paid to his aged mother." The collection will be a welcome resource for students, teachers, and researchers. —MB



Space Atlas

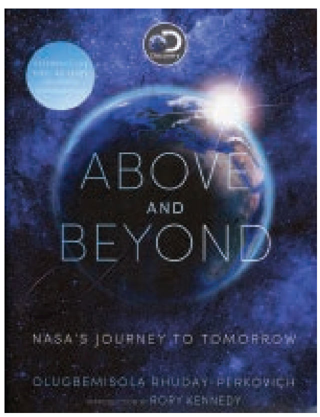
Mapping the Universe and Beyond

James Trefil
National Geographic, 2018 (2nd ed.). $50.00

In anticipation of the 50th anniversary of the first Moon landing, National Geographic has revised its *Space Atlas*, a beautifully illustrated compendium of photos, graphics, maps of our solar system, and more, with text written by physicist James Trefil. Because of the many discoveries made since the atlas was first published in 2012, new material has been added on topics such as gravitational waves, Mercury's polar craters, and the dwarf planet Ceres. Former astronaut Buzz Aldrin, the second person to walk on the Moon, provides a new foreword in which he discusses Earth's natural satellite and its role in space exploration. —CC

Above and Beyond

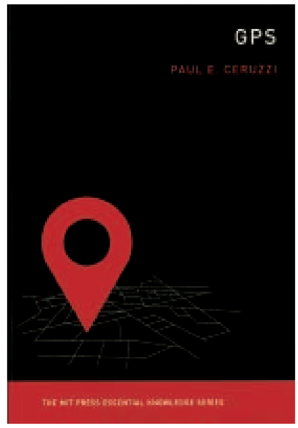
NASA's Journey to Tomorrow



Olugbemisola Rhuday-Perkovich
Feiwel and Friends, 2018. $19.99

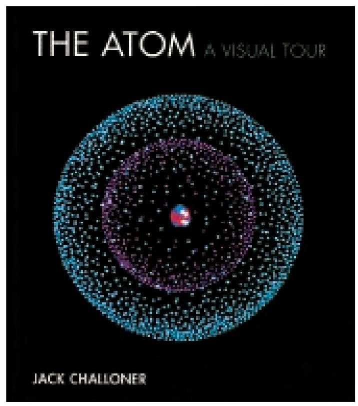
To "look back over the last sixty years and look forward to the next" is the goal of this glossy coffee-table book filled with photos and illustrations of spacecraft and space phenomena, profiles of significant figures in NASA's history, and fun facts and trivia. Published to accompany a documentary film of the same name, *Above and Beyond* starts with the creation of NASA in 1958 and covers the agency's original goals, progress over the past six decades, and plans for the future. Although the book is a celebration of all things NASA, author Olugbemisola Rhuday-Perkovich does touch on such issues as racial and gender discrimination in NASA's early history. —CC

GPS



Paul E. Ceruzzi
MIT Press, 2018. $15.95 (paper)

Part of the MIT Press Essential Knowledge series, *GPS* centers on the history and development of the satellite-based radio-navigation system that has become ubiquitous. Author Paul Ceruzzi, curator at the Smithsonian Institution's National Air and Space Museum, starts off with a brief overview of early navigation systems, then launches into a discussion of how and why GPS developed and became the military and commercial behemoth it is today. The pocket-sized volume is illustrated with black-and-white photos and includes a timeline and suggested further reading. —CC



The Atom

A Visual Tour

Jack Challoner
MIT Press, 2018. $33.00

Beautifully illustrated with photos, diagrams, and artistic renderings, *The Atom: A Visual Tour* lives up to its title. The excellent imagery complements science writer Jack Challoner's text, in which he presents a comprehensive study of what was once thought to be the fundamental building block of matter. As Challoner shows, our concept of the atom has evolved over the past 2500 years as we've learned more about its structure and properties through our ever-improving technologies. Aimed at the general reader, the text is an introduction to current theory and what it has revealed about our world. —CC PT

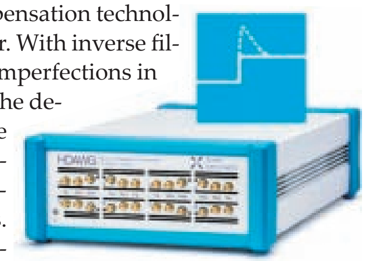
NEW PRODUCTS

Focus on test, measurement, software, and instrumentation

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

Andreas Mandelis

Arbitrary waveform generator with precompensation



Zurich Instruments now offers real-time precompensation technology for its HDAWG arbitrary waveform generator. With inverse filtering, precompensation minimizes the effect of imperfections in the wiring and ensures that the signal applied to the device under test equals the signal designed on the HDAWG. Available filters can correct for AC coupling, spurious inductance and capacitance, impedance mismatch, amplifier ringing, and other effects. Real-time precompensation can be applied individually to each channel. It needs to be configured only once for each wiring arrangement; testing and simulation tools are provided to facilitate configuration. The HDAWG precompensation will benefit quantum computing applications that use flux bias and gate voltage pulses. Other applications include electron paramagnetic resonance and nuclear magnetic resonance. ***Zurich Instruments AG***, *Technoparkstrasse 1, 8005 Zürich, Switzerland, www.zhinst.com*

Low-pulsation diaphragm pump

The FP 400 diaphragm pump from KNF combines the traditional advantages of diaphragm-pump technology—it is self-priming, can run dry, and has a long operating lifetime—with a pulsation level comparable to those of gear pumps. The FP 400 offers pulsation less than 150 mbar; lower levels can be achieved, depending on factors such as flow-path length and configuration. It delivers up to 5 L/min of liquid at back pressures to 15 psi. The flow is fully stable with fluid viscosities between 1 cSt and 150 cSt. Viscosities up to 500 cSt can be handled with some flow-rate reduction. The FP 400 provides gentle, low-shear conveyance of sensitive media. It features very low vibration, a noise level below 55 dBA, an IP65 protection rating, and chemically resistant flow-path-material options for use with aggressive media. It is suitable for recirculation applications in various fields, including semiconductors and fuel cells. ***KNF Neuberger Inc***, *2 Black Forest Rd, Trenton, NJ 08691-1810, www.knfusa.com*

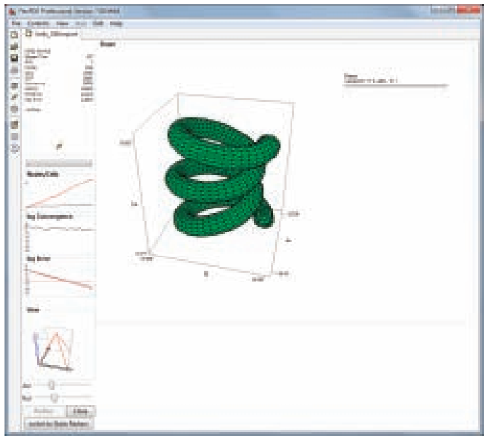
High-speed amplifier

Aerotech's XL4s network digital drive is a high-performance linear amplifier designed to provide closed-loop servo control of voice-coil and single-phase motors and eliminate the nonlinearities common with pulse width-modulation amplifiers. According to the company, the XL4s outperforms other single-phase motor controllers because of its high 192 kHz servo rate. That correlates to both better tracking of errors and better quality of parts at high speeds in applications such as fast-tool servos and high-dynamic optical focusing axes. The amplifier offers a fiber-optic interface and easy software setup. A floating-point digital signal processor controls the proportional-integral-derivative functions. Aerotech claims that advanced features such as full state feedforward control and look-ahead-based velocity control help XL4s users achieve low settling times, long-term thermal stability, and submicron-level tracking accuracy. ***Aerotech Inc***, *101 Zeta Dr, Pittsburgh, PA 15238, www.aerotech.com*

Wideband vector signal generators

In a single test instrument, each of the two signal generators in Keysight's VXG series delivers an optimized 5G New Radio test system setup that leverages dual-channel 44 GHz vector signal generation with up to 2 GHz RF modulation bandwidth and phase-coherent capability. The microwave signal generators address demanding wideband millimeter-wave applications for 5G and satellite communications. They decrease test setup complexity and reduce path losses introduced in over-the-air test environments. The VXG series is offered in benchtop and modular form factors. Their low phase noise and distortion, high output power, and excellent modulation make them suitable signal generators for a wide range of applications in wireless communications and aerospace defense industries. ***Keysight Technologies Inc***, *1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*

Scientific modeling software

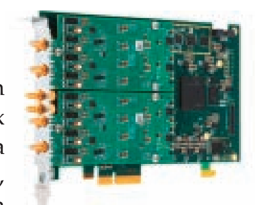


PDE Solutions has introduced the latest version of its FlexPDE Multiphysics software, which provides computational support for solving partial-differential-equation systems in science and engineering. Revisions in version 7 include the removal of diagnostic blocking staged eigenvalue runs and the addition of a diagnostic for duplicate region definitions. The software now correctly suppresses prompts when using the −S switch in the nongraphical user interface version and changes the default values for NGRID in 2D and 3D. Corrections have been made to an occasional gridder cross-link in periodic domains and to some errors, such as the one that was in the table-bounds report and the one that was causing erratic behavior in the convergence of linear steady-state equations. ***PDE Solutions Inc****, 9408 E Holman Rd, Spokane Valley, WA 99206, www.pdesolutions.com*

Miniature linear stages

The L-505 miniature-linear-stage series from Physik Instrumente provides high-precision motion in a compact, economical package. The series includes various drive and configuration options, from open-loop stepper motors with lead screws to fast, servo-motor-driven units with linear encoders and low-friction ball screws. Two basic designs are offered: a shorter, 60-mm-wide, lower-profile 21 mm version with a folded drive train and the motor side by side with the platform; and a longer, 36-mm-wide, higher-profile 25 mm inline version. Both come in travel ranges of 13 mm and 26 mm. The *XY* combinations can be assembled without adapter plates; a *Z* bracket is available for *XYZ* assemblies. For the most demanding applications, such as the fast, precise alignment of photonics components, an *XYZ* piezo scanner option can be added to the linear stages. ***Physik Instrumente LP****, 16 Albert St, Auburn, MA 01501, www.pi-usa.us*

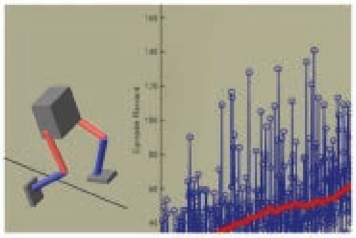
Ultrahigh-frequency AWGs



Spectrum Instrumentation has released six arbitrary waveform generators optimized for signal quality, size, and cost. The M2p.65xx series offers the latest 16-bit digital-to-analog converters (DACs), a fast PCIe x4 interface with a streaming speed up to 700 megabytes/s, and a card length of only 168 mm to fit into nearly every PC. With a speed of 40 megasamples/s or 125 megasamples/s, high onboard memory of 512 megasamples, output levels of up to ±6 V, and four additional multipurpose outputs, the cards are suitable for use in signal generators operating at frequencies between 1 MHz and 60 MHz. Applications include ultrasound, laser, lidar, radar, medical science, and big physics experiments. Models have one, two, or four channels per card, and each channel has its own DAC and output stage. Multichannel cards share a common clock and trigger to ensure full synchronization. ***Spectrum Instrumentation Corp****, 15 Warren St, Ste 25, Hackensack, NJ 07601, https://spectrum-instrumentation.com*

Multisequence AWGs

Tektronix has launched its AWG70000B series of arbitrary waveform generators to support advanced research and other applications that require the ability to dynamically alter signal sequences during test scenarios. The new AWGs feature waveform memory of 32 gigasamples and the company's Streaming Waveform ID functionality, which provides users with immediate access to a total of 16 383 sequence steps through a direct Ethernet interface. The new capabilities allow the AWG70000B to replicate the chaos of the real world during evaluation of modulated-signal formats and electronic-warfare-simulation exercises. For example, in wireless communications research, users can change modulation types to simulate Doppler radars, building obstructions, or other obstacles to improve the durability of orthogonal frequency-division multiplexing signals. ***Tektronix Inc****, 14150 SW Karl Braun Dr, PO Box 500, Beaverton, OR 97077, www.tek.com*



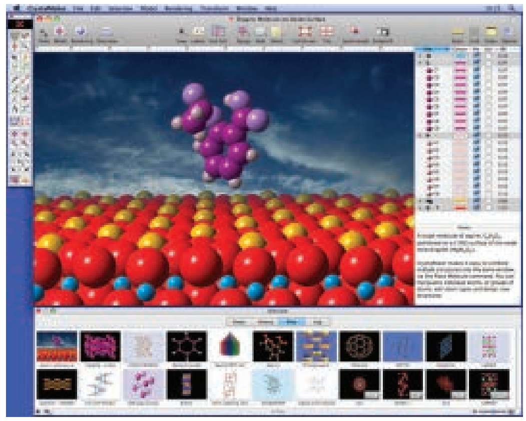
Mathematical programming software

Release 2019a of MATLAB and Simulink from MathWorks contains new products and enhancements for artificial intelligence (AI), signal processing, and static analysis and new capabilities and bug fixes across all product families. The reinforcement learning toolbox improves the MATLAB workflow for AI by facilitating a type of machine learning that trains an "agent" through repeated trial-and-error interactions with an environment to solve controls and decision-making problems. Enhancements to the computer-vision and data- and image-acquisition toolboxes further support AI. New signal-processing and communications products promote wireless and electronics development. For example, a Simulink add-on, Mixed-Signal Blockset, provides fast model construction, rapid simulation, and deep insights into mixed-signal system design models with dedicated analysis and visualization tools. ***The MathWorks Inc****, 1 Apple Hill Dr, Natick, MA 01760-2098, www.mathworks.com*

Optical tensiometer

The Attension Theta Flex contact angle meter from Biolin Scientific has a modular design with several measurement options and all-inclusive software. It features advanced imaging and analysis algorithms to detect and precisely gauge static and dynamic contact angles and surface free energy. The effect of roughness to wettability can be measured with the new 3D topography module. The Theta Flex also analyzes surface and interfacial tension and interfacial rheology. All steps from loading and performing the measurement to analyzing the data can be automated, and disposable liquid tips eliminate the need for preparations and cleaning. The tensiometer is suitable for research and quality control in various research and industrial applications, including chemicals, pharmaceuticals, electronics, and energy. ***Biolin Scientific Inc****, 514 Progress Dr, Linthicum Heights, MD 21090, www.biolinscientific.com*

Crystal structure software



CrystalMaker has made available the latest version of its software for visualizing crystal and molecular structures in research and teaching in chemistry, solid-state physics, materials science, mineralogy, and crystallography. Version 10.4 includes more than 60 new features and 100 new structures, with over 70 new minerals. The structures library window has been redesigned for easy access to documentation and incorporates CrystalViewer for visualizing structures. A new modeling engine takes torsion angles into account and automatically detects rings. The software offers a dark mode for Mac users, a live powder-diffraction mode for physicists, an interpolate structures command for crystallographers, and customizable axes. The company has also updated its CrystalDiffract powder-diffraction software to version 6.8. ***CrystalMaker Software Ltd****, Centre for Innovation & Enterprise, Oxford University Begbroke Science Park, Woodstock Rd, Begbroke, Oxfordshire, OX5 1PF, UK, http://crystalmaker.com*

High-speed floating picoammeter

A new version of RBD Instruments' 9103 USB picoammeter features more reads per second and 5000 DC volts of isolation to chassis ground. According to the company, it may open up research possibilities such as DC measurement of very small electron and photomultiplier signals. Electron- and ion-beam measurements can be biased to reduce secondary electrons or retard the beam as needed for experiments. Designed to provide accurate current measurements in noisy environments such as synchrotron beamlines, the 9103 can capture bipolar DC from low picoamps to milliamps. The new option increases the number of reads per second from 40 to more than 500, which is fast enough to perform optical chopper experiments and improves accuracy. Because the 9103 can be synced, users can configure a multichannel DC picoammeter to have up to 256 channels and high speed, high voltage, or both. ***RBD Instruments Inc****, 2437 NE Twin Knolls Dr, Ste 2, Bend, OR 97701, https://rbdinstruments.com*

Laser frequency phase-lock stabilizer

The Moku:Lab Laser Lock Box from Liquid Instruments uses high-performance modulation locking techniques to stabilize a laser's frequency to a reference cavity or atomic transition. The instrument can generate and output modulation signals at up to 200 MHz and demodulate the input signal with an internal or external local oscillator, including a phase-locked loop option. It can scan across resonances with sawtooth or triangle waveforms at up to 1 MHz, quickly lock to any zero crossing in the error signal by using the tap-to-lock feature, filter the demodulated input signal with up to fourth-order infinite-impulse response filters, and individually configure fast and slow proportional-integral-derivative controllers. With the instrument's integrated two-channel oscilloscope, users can observe signals at any point in the signal-processing chain at up to 500 megasamples/s. ***Liquid Instruments***, *2223 Avenida De La Playa, Ste 204, La Jolla, CA 92037, www.liquidinstruments.com*

Electro-optical software

Electro Optical Industries has announced its latest Infratest electro-optical software for research and industrial testing of visible, near-IR, and IR cameras; intensified CCDs; goggles; and laser range finders. The software allows parameters such as distortion to be accurately measured with a mirror-based collimator, which enables determination of resolution and ranges at the edge of the wide-field-of-view camera. It is now possible to remotely monitor the temperature of blackbodies' cooling fluid. Infratest calculates key functions such as minimum resolvable temperature difference by automatically taking into account the background temperature data simulated by a secondary background blackbody. ***Electro Optical Industries Inc***, *320 Storke Rd, Ste 100, Goleta, CA 93117, www.electro-optical.com*

Capacitance manometer

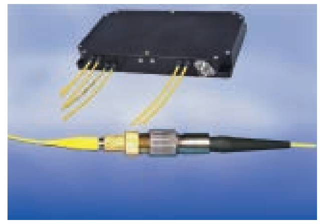
According to the Kurt J. Lesker Co, its KJLC Carbon is the most cost-effective capacitance manometer on the market, yet it maintains a high accuracy rate and a fast response time. Conceived for research use, the gauge uses an ultrapure alumina ceramic diaphragm, which is corrosion-proof and allows for better signal stability. The Carbon offers direct measurement of chamber total pressure independent of gas type or composition and eliminates lookup tables and conversion factors. Long-term output stability makes possible state-of-the-art process repeatability. A compact, simplified design allows for placement in space-restricted areas. The Carbon can be mounted in any orientation. ***Kurt J. Lesker Company***, *1925 Rte 51, Jefferson Hills, PA 15025, www.lesker.com* PT

PHYSICS TODAY

PRODUCT PICKS

OCT COMPONENTS

OZ has introduced a line of OCT Fiber Optic Components: High Speed Polarization Controller/Scrambler, Electrically/Manual Controlled Optical Delay Line, Collimators/Focusers, Attenuator, Faraday Rotators/Mirror, Isolator, Ultra Stable Laser Module, Fused Coupler, Polarizers, Reflector, Power Monitor, Circulator, Tunable/Inline Fabry-Perot Filter



OZ Optics Ltd
www.ozoptics.com

Helmholtz Coil Systems

- System with Active Cancellation available
- 350mm to 2m diameter coils
- Maximum field strength from 250µT to 1mT at DC

US distributor:
GMW Associates
gmw.com • bartington.com
Bartington™ Instruments

ASSOCIATE/FULL PROFESSOR POSITION IN COMPUTATIONAL CONDENSED MATTER PHYSICS THE UNIVERSITY OF TENNESSEE

The Department of Physics and Astronomy in collaboration with the Department of Electrical Engineering and Computer Science (EECS) at The University of Tennessee (UT) invites applications to fill a tenured faculty position at the associate/full professor level. The successful candidate will hold a joint appointment in the physics (primary) and EECS departments. This search aims to strengthen our efforts in the development of new algorithms for computational condensed matter physics research. A further expansion in the area of quantum materials and quantum information is anticipated, which will be supported by junior-level hires following the successful completion of this search.

Candidates should have a PhD in Physics or related field, a strong research record in computational condensed matter physics, experience in the development of simulation algorithms for quantum materials, and background in computer science with an emphasis on novel approaches such as machine learning and other promising techniques. The candidate is expected to provide leadership in developing a synergistic interdisciplinary quantum materials program, establish an externally funded research program, provide interdisciplinary training for graduate students and postdoctoral researchers, and to contribute to the teaching mission of the departments. While the preferred expertise should be in the broad area of algorithmic development for quantum many body physics, a strong interest in bridging the efforts of the above-mentioned departments is highly desirable.

UT Knoxville is Tennessee's flagship state research institution. The successful candidate will benefit greatly from available computational resources and by the proximity to research facilities at Oak Ridge National Laboratory, including the Joint Institutes for Computational Sciences, Advanced Materials, and Neutron Sciences.

UT Knoxville is seeking candidates who have the ability to contribute in meaningful ways to the diversity and intercultural goals of the University. Applicants should send a CV, list of publications, a description of research and teaching experience, a proposed research program, and arrange for at least three letters of reference to be submitted separately. All application materials, including the letters, should be submitted via email to https://apply.interfolio.com/61568 We will start reviewing applications by June 1, 2019.

The University of Tennessee is an EEO/AA/Title VI/Title IX/Section 504/ADA/ADEA institution in the provision of its education and employment programs and services. All qualified applicants will receive equal consideration for employment and admission without regard to race, color, national origin, religion, sex, pregnancy, marital status, sexual orientation, gender identity, age, physical or mental disability, genetic information, veteran status, and parental status.

OBITUARIES

Manfred Eigen

German Nobel laureate Manfred Eigen, who made towering contributions to chemical physics, biophysics, and molecular evolution, died in Göttingen, Germany, on 6 February 2019.

Eigen was born on 9 May 1927 in Bochum, Germany. His formal studies were interrupted by World War II. In 1942, at age 15, he was drafted into the German air force. Toward the end of the war, he was taken prisoner by the Allies but managed to escape and return home on foot, walking nearly 1000 km. He resumed his education at the University of Göttingen when the institution reopened its doors after the war. Eigen enrolled in geophysics, not his first choice but the closest match given the scarcity of available slots. Werner Heisenberg was one of his professors. Eigen earned a doctorate in 1951 under the tutelage of Arnold Eucken with a dissertation on the specific heat of aqueous electrolyte solutions and heavy water.

Eigen then took a research post at the Max Planck Institute for Physical Chemistry, which had recently been established in Göttingen. He conducted research into such topics as proton-transfer reactions in ice crystals, thermal conductivity, sound absorption, and reactions of metal-ion complexes. He also learned that certain chemical reactions were deemed "immeasurably fast," a description he found particularly disturbing. Unsatisfied, he took up the challenge of determining the rates of such elusive processes, an effort that proved most rewarding.

In the early 1960s, Eigen's career took a meteoric rise when he demonstrated how to measure ultrafast chemical reactions that occur on submicrosecond and even nanosecond time scales. Those reactions were deemed unmeasurable because the time scales were shorter than the times required for the reactants to be fully mixed. Particularly opaque were proton-transfer reactions in aqueous media, which are ubiquitous in biochemical processes. In a brilliant turn, Eigen let the reaction reach equilibrium, perturbed that state with an ultrafast sonic or light pulse, and spectroscopically monitored the system's relaxation back to equilibrium. The relaxation parameters provided the necessary information to obtain the reaction rates and even provided mechanistic insights on how the reactions took place.

In 1964 Eigen presented his research at the Faraday Society in London and achieved instant fame as one of the foremost experimentalists of his day. That same year he became head of the Max Planck Institute in Göttingen, which would later become the Max Planck Institute for Biophysical Chemistry. In 1967 Eigen shared the Nobel Prize in Chemistry with Ronald Norrish and George Porter for his work on ultrafast reaction measurements.

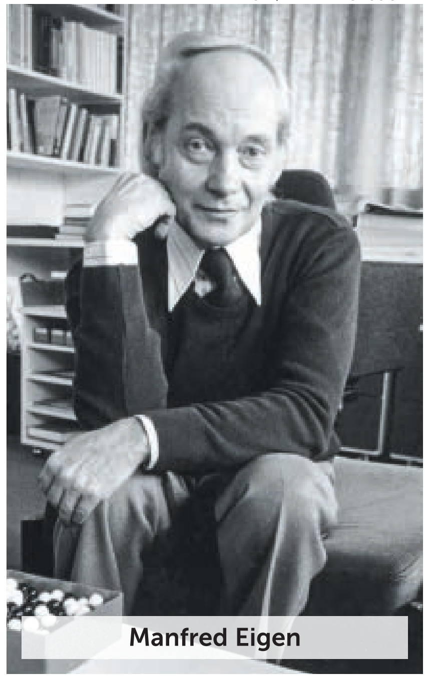
Eigen's research in the 1970s took a different direction: He became more involved in the field of molecular evolution. He built a kinetic scheme known as a hypercycle that he believed would capture the essential features of self-organization in prebiotic systems at molecular scales and would contribute to a cogent explanation for the emergence of biological information. Novel concepts like quasispecies arose from such forays.

In 1971 Eigen posed an information-theoretic paradox that still stands today: Without editing enzymes, the size of a replicating molecule is limited, since otherwise an error catastrophe would arise from accrued mutations over generations. Yet for a replicating molecule to encode editing enzymes, it must be substantially larger than that limit. That line of work proved inspirational to a generation of researchers, including 2018 Nobel laureate Frances Arnold, who tested Darwinian scenarios at the molecular level. It also fostered multidisciplinary efforts that became the hallmark of the Max Planck Institute under Eigen's leadership.

With a Nobel Prize to his name, Eigen became an icon of postwar German science and a key player in the restoration of its former glory. Predictably, he was a commanding figure in the German scientific establishment. In 1971 the Max Planck Society created—essentially for him—the Institute for Biophysical Chemistry. Although he personally supervised its foundational work, Eigen eventually turned down the post of permanent director and instead headed the institute's department of biochemical kinetics until his retirement in 1995.

Eigen was a great communicator. Akin to his beloved Wolfgang Amadeus Mozart, Eigen had a style marked by clarity and rigor. He published three books aimed at a general audience: *Laws of the Game: How the Principles of Nature Govern Chance* (1981), *Steps Towards Life: A Perspective on Evolution* (1992), and *From Strange Simplicity to Complex Familiarity: A Treatise on Matter, Information, Life and Thought* (2013). He and his scientific partner Ruthild Winkler-Oswatitsch, who eventually became his second wife, collaborated on those popular accounts.

FILSER/MAX PLANCK SOCIETY



Manfred Eigen

Eigen was inspirational and humorous, with a sense of irony. As a former senior researcher in his division at the Max Planck Institute, I recall once giving a complicated mathematical presentation on a scenario for the origin of biological information, with Eigen and his group in full attendance. As my derivations were getting more and more convoluted, Eigen politely interrupted me and said, "Ariel, if you want to go to Stockholm, never get past the linear approximation."

Ariel Fernández
Bahia Blanca, Argentina PT

INDEX TO ADVERTISERS

COMPANY	PAGE NO.
Aerotech Inc	24
Amptek Inc	62
Applied Physics Reviews	33
Bartington Instruments	66
Cobolt AB	32
COMSOL Inc	C3
Cremat	30
Cryogenic Control Systems Inc	65
Duniway Stockroom Corp	59
Electro-Optical Products Corp	23
Hiden Analytical Inc	21
Hinds Instruments Inc	59
Inrad Optics	13
Janis Research LLC	18
Mad City Labs Inc	12
Mathworks Inc	C4
Nanomagnetics Instruments	1
National Reconnaissance Office	9
OriginLab Corporation	15
Oxford Instruments	7
OZ Optics Ltd	66
PDE Solutions Inc	46
Pearl Companies	C2
Spectrum Instrumentation GmbH	64
Stanford Research Systems	3, 25
Toptica	63
UC Components	28
University of Tennessee	66
Zurich Instruments AG	4

ADVERTISING SALES MANAGEMENT

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

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

Recruitment Advertising
ptjobs@aip.org

Ad Traffic Coordinator
Unique Carter • ucarter@aip.org

ADVERTISING SALES CONSULTANTS

USA and CANADA

AL, AR, AZ, CO, DC, FL, IA, ID, IL, IN, KS, KY, LA, MD, MI, MN, MO, MS, MT, ND, NE, NM, NV, OH, OK, SD, TN, TX, UT, VA, WI, WV, WY, and West Canada
Mike Shevlin
p: +1 847.749.0168
m: +1 847.602.6844
mshevlin@theshevlingroup.com

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

Merrie Lynch
+1 617.357.8190
merrie.lynch@celassociates2.com

INTERNATIONAL

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

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

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

PHYSICS TODAY

www.physicstoday.org/advertising

QUICK STUDY

Howard Milchberg is a professor of physics and electrical engineering at the University of Maryland in College Park.

Indestructible plasma optics

Howard Milchberg

Working with ultra-intense laser pulses? These optical elements are what you need.

The everyday concept of optics is of transparent elements, such as glass lenses, that bend beams of light in useful ways. The small lenses in our smartphones are now almost as ubiquitous as the lenses in our eyes. In both cases, the lenses redirect the rays of light scattered from, say, a tree and project them to form an image of the tree on the camera's photosensitive chip or on our retinas.

Suppose you directed a laser beam into your smartphone lens. (Don't even think about doing the same with your eye.) The lens redirects the beam to a near-point-like focal spot on the chip. The milliwatts of power delivered by common laser pointers is more than enough to damage your smartphone. But what if you dialed up the beam's power enough that the beam damaged the lens before arriving at the focus? For a high-average-power continuous-wave beam, the small fractional optical absorption that always takes place inside transparent dielectric materials would eventually heat and thermally stress the lens until it fractures and melts. The lens would be ruined.

Another type of beam, though, has a radically different effect on the lens: That beam is an ultrahigh-peak-power laser pulse formed by packing a modest amount of energy into an extremely short-duration pulse. Half of the 2018 Nobel Prize in Physics was awarded for precisely that feat of compression (see PHYSICS TODAY, December 2018, page 18). If such a now-routine pulse—typically of a peak intensity up to 10^{22} W/cm^2 and a duration shorter than 100 fs—is incident on the lens, the laser electric field would cause electrons to nearly instantaneously tunnel out of the bound states of surface atoms. The laser-induced tunneling would form a solid-density plasma with optical properties akin to a highly polished metal mirror, and the pulse would specularly reflect from the surface.

To generate the plasma, one would need to focus the beam on the surface, and the interaction would need to take place in vacuum to prevent the laser ionization of air that would defocus the pulse well before it arrived at the surface. Long after the pulse is gone, damage follows on a nanosecond acoustic time scale as the dense hot plasma (with temperature on the order of 10^6 K and a pressure of 10^7 atmospheres) launches an impulsive pressure spike into the bulk of the glass and causes significant local damage.

Plasma mirror

I have just described the simplest kind of plasma optic: a plasma mirror. And I wouldn't waste a lens for the job. A glass slide would work fine; it could be translated to a fresh location after every laser pulse. The first experiment to do such a thing simply aimed to understand the optical properties of a hot, solid-density plasma by measuring its reflection of ultrashort, intense pulses.

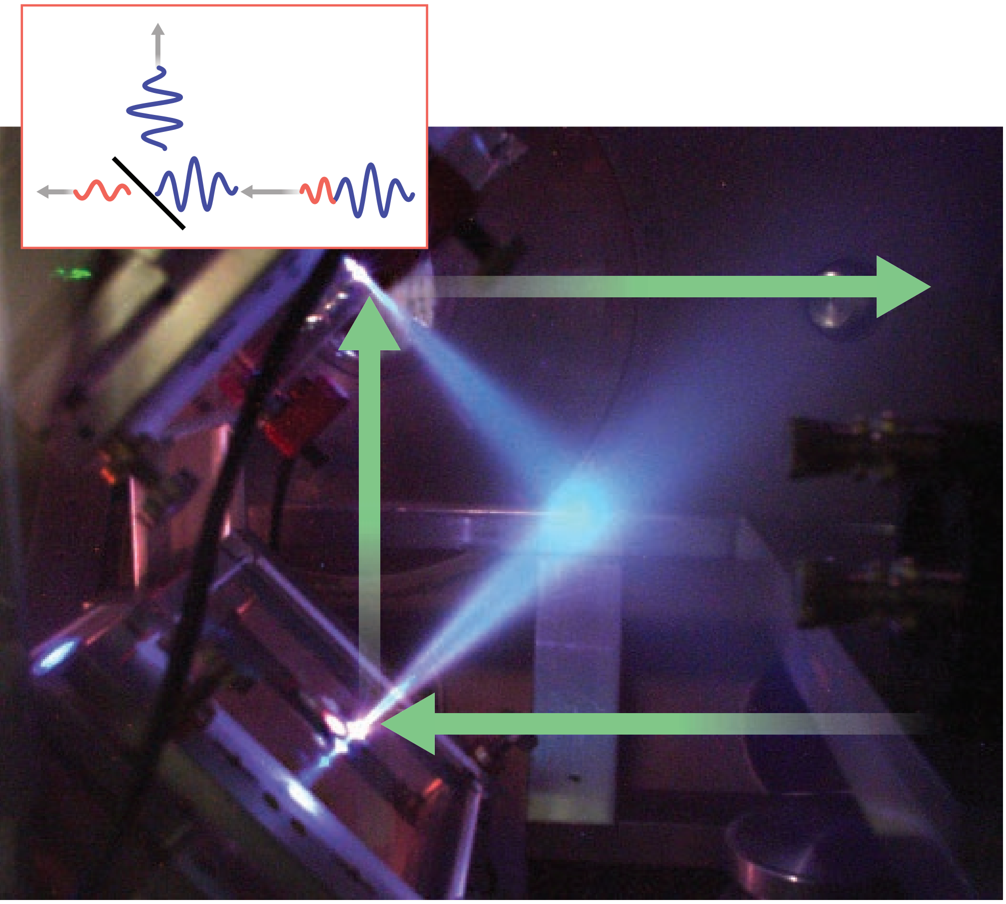
Plasma mirrors were later used to sharpen the leading edge of short pulses and eliminate an undesirable feature known as "prepulse"—laser energy that precedes the steeply rising main pulse. Tuning the focal spot size at the surface can ensure that the early, lower-intensity portion of the pulse is below the threshold for ionization and transmits through the glass, whereas the intense portion ionizes the surface and reflects from its self-generated plasma mirror, as shown in the figure. Having a prepulse-free laser pulse is essential to several high-intensity laser–matter interaction experiments, including the laser-driven acceleration of protons, widely pursued as a potential radiation therapy. (See the article by Jerimy Polf and Katia Parodi, PHYSICS TODAY, October 2015, page 28.)

In the plasma mirror, the intense pulse is reflected from the surface rather than transmitted, because the plasma electron density exceeds the wavelength-dependent critical density. You can see the true pallor of your face in the morning (assuming you want to) because the electron density of the aluminum layer on your bathroom mirror, at roughly 10^{23} cm^{-3}, greatly exceeds the critical density and ensures reflection for all wavelengths in the visible spectrum. At densities above critical—for hot plasma mirrors and for bathroom mirrors—the incident wave undergoes a phase shift that favors reflection over forward propagation, or transmission. To understand that behavior, note that at densities above critical (or, equivalently, for laser frequencies below the plasma frequency) the plasma electrons can follow the swings of the laser electric field and short it out. When the wave is "shorted out," its response is to reflect rather than transmit.

Gratings, fibers, and waveplates

What about other common optical elements, such as curved mirrors, gratings, optical fibers, and waveplates? They all have plasma analogues. The curved mirror is straightforward: The plasma generated on a curved surface can redirect an intense pulse to converge or diverge, and it can serve as both a prepulse filter and a focusing element for proton acceleration. As for gratings, intersecting two laser beams in a gas or subcritical-density plasma generates an interference pattern, which is a volume distribution of alternating bright and dark fringes. If the beams are intense, the bright fringes generate localized ionization and form a volume plasma grating.

The grating can diffract each beam in the direction of the other and, depending on the beams' polarizations and wavelengths, acts as a controllable birefringent element, or wave-



A DOUBLE PLASMA MIRROR in action. A focused 50 fs, 3 J titanium–sapphire laser pulse reflects from self-generated plasma mirrors on two 10 cm x 10 cm quartz plates in an experiment at the École Polytechnique in Paris. The two mirrors eliminate the beam's prepulse—its temporal leading edge (red in the inset, its amplitude greatly exaggerated for clarity)—before the main pulse (blue) is recollimated for experiments downstream. The intense part of the pulse forms the plasma mirror it reflects from. The photograph shows plumes of plasma ejected normal to the mirror surfaces; the plumes accompany impulsive shock waves launched into the bulk of the quartz, long after the pulse has gone. The green arrows show the beam path. (Adapted from R. Marjoribanks et al., in *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science and Photonic Applications Systems Technologies*, Optical Society of America, 2005, paper JFA7.)

plate, on one or both beams. Because of the diffractive scattering, plasma gratings can also mediate energy transfer from one beam to another and are used to balance the laser flux across multiple high-energy beams focused inside fusion targets at the National Ignition Facility. Plasma gratings can even be used to make holograms: Object-encoded two-beam interference patterns can be imprinted on solid surfaces; a third intense beam reflects from the patterned plasma mirror to read the hologram.

An essential use of plasma optics is in laser-driven electron acceleration, which requires propagation of extremely intense pulses through long distances $L \gg z_0$ of subcritical-density plasma, where the Rayleigh length z_0 is the characteristic distance over which the beam's intensity would normally drop by a factor of two due to diffractive divergence.

At low laser intensities, glass-fiber waveguides defeat beam divergence and form the backbone of today's global communications systems. In a glass fiber, the refractive index is bigger on axis than off, which makes the light wave slowest on axis and curves its phase fronts inward, canceling the outward curvature from diffraction. Unfortunately, the high intensities—greater than 10^{18} W/cm^2—needed for laser-based accelerators are at least six orders of magnitude beyond the glass fibers' destruction threshold.

An optical fiber made of subcritical-density plasma would work quite nicely as a waveguide. Because the free electrons of a plasma reduce the refractive index, the fiber's on-axis density must be lower than off axis. Two techniques do the job, and future accelerators will likely depend on them: the electric-discharge capillary waveguide and the laser-driven hydrodynamic plasma fiber.

In the electric-discharge capillary, the desired waveguiding profile results from plasma cooling near the capillary inner wall, which increases the density there. In the plasma fiber, a long, 10-μm-diameter, hot plasma is generated in a gas by a short laser pulse. The hot, thin plasma cylinder explodes radially into the background gas and creates an electron-density minimum on axis with high outer walls. That is exactly the refractive-index profile suitable for the injection and guiding of a separate high-intensity pulse that accelerates electrons. Although the cylindrical explosion is transient, the guide appears stationary to an intense speed-of-light pulse propagating along it.

Although plasma optics are seemingly highly exotic, they may soon become routine. Laser technology has advanced to a stage where the widely available, compact, short-pulse lasers used for machining and materials processing can also induce plasma optics, where even relativistic effects are important. No longer an exotic application, plasma optics will become a tool in the optical toolbox.

Additional resources

▶ H. M. Milchberg et al., "Resistivity of a simple metal from room temperature to 10^6 K," *Phys. Rev. Lett.* **61**, 2364 (1988).
▶ P. Michel et al., "Dynamic control of the polarization of intense laser beams via optical wave mixing in plasmas," *Phys. Rev. Lett.* **113**, 205001 (2014).
▶ S. H. Glenzer et al., "Symmetric inertial confinement fusion implosions at ultra-high laser energies," *Science* **327**, 1228 (2010).
▶ A. Leblanc et al., "Plasma holograms for ultrahigh-intensity optics," *Nat. Phys.* **13**, 440 (2017).
▶ C. G. Durfee III, H. M. Milchberg, "Light pipe for high intensity laser pulses," *Phys. Rev. Lett.* **71**, 2409 (1993).
▶ Y. Ehrlich et al., "Guiding of high intensity laser pulses in straight and curved plasma channel experiments," *Phys. Rev. Lett.* **77**, 4186 (1996). PT

Spore vortex

The spreading of plant and fungal spores is of great importance to plant and animal health. Once lofted by the wind, the particles can travel anywhere from a few kilometers to thousands and can even cross continents. Rainfall can drive the dispersal. When raindrops land on a spore-laden leaf, they spawn droplets that envelop some of the particles and carry them short distances. Cornell University's Sunghwan Jung and colleagues have now shown that raindrop-triggered transport is dominated by a second mechanism: dry-spore dispersal.

This image, created from several superposed high-speed video frames, captures raindrop-induced trajectories of glass beads, which the researchers used as surrogates for actual spores. After a raindrop hits the leaf surface, it starts to spread, and spores that get trapped at the advancing meniscus will collide with and eject other dry spores. The raindrop impact also generates an air vortex, which can lift the ejected spores several millimeters, enough for them to be picked up by the wind and carried away over long distances. (S. Kim et al., *Proc. Natl. Acad. Sci. USA* **116**, 4917, 2019. Image by Seungho Kim, Hyunggon Park, and Sunny Jung, submitted by Sunny Jung.) —RJF