PHYSICS TODAY

December 2020 • volume 73, number 12
A publication of the American Institute of Physics

SOUND affects

Making robots microscopic

Hydrogen-powered aircraft

Boosting radiation therapy

APSIT
INSURANCE FOR SCIENCE PROFESSIONALS

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

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

DID YOU KNOW?

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

APPLY ONLINE TODAY!

APSITPLANS.COM/LTL-NOW | 800.272.1637

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

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

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

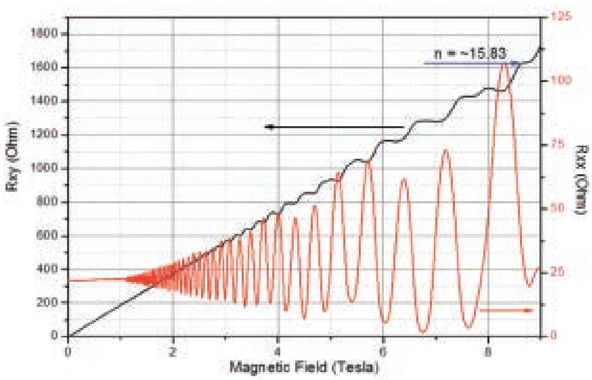
192136-APSIT-LTL-MAG-PAD

NanoMagnetics Instruments

HEMS

Technical Specifications

- 3-1,500K temperature range
- Optical access
- Multi-sample experiments with 4-contact van der Pauw and 6-contact Hall
- Wide range of materials; GaAs, InP, InAs, Si, Ge, SiGe, HgCdTe, GaN, SiC, AIN, metal oxides and organic conductors
- Ideally suited for materials research, product development and quality control

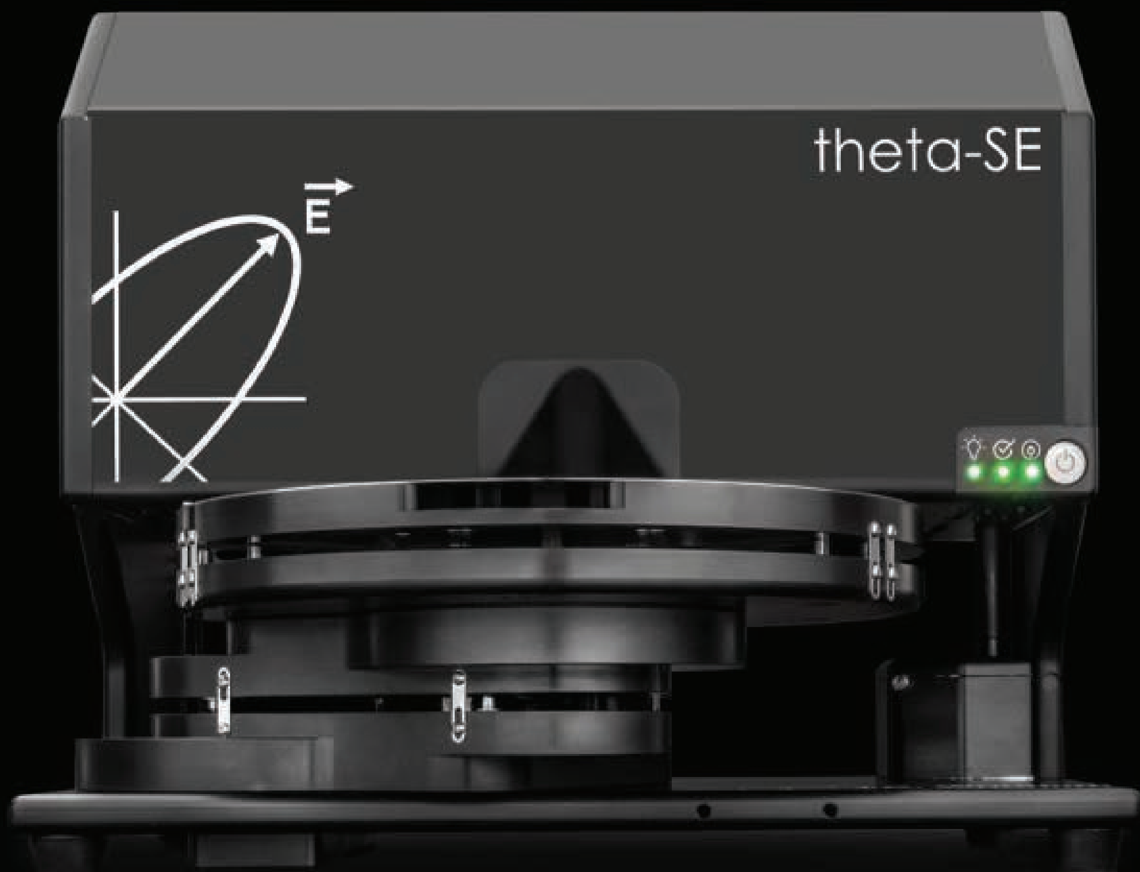


Shubnikov de Haas Oscillations in GaInAs Quantum Well

sales@nanomagnetics-inst.com
www.nanomagnetics-inst.com
+44 7906 159 508
Suite 290, 266 Banbury Road
Oxford OX2 7DL, United Kingdom
NMInstruments

Driving Innovation in Spectroscopic Ellipsometry

Introducing the new, fully-integrated theta-SE



ACCURATE. COMPACT. HIGH SPEED.

+Characterization of thin film refractive index & thickness

+Uniformity mapping of samples as large as 300 mm

+Full spectral measurement in less than one second

J.A. Woollam

The next generation ***Lock-In Amplifiers***
Only from SRS !

The SR86x series brings new performance to lock-in measurements — a frequency range of 4 MHz (SR865A) or 500 kHz (SR860), state-of-the-art current and voltage input preamplifiers, a differential sinewave output with DC offset, and fast time constants (1 μs) with advanced filtering.

And there's a colorful touchscreen display and a long list of new features ...

- ☑ Deep memory data recordings
- ☑ FFT analysis
- ☑ Built-in frequency, amplitude & offset sweeps
- ☑ 10 MHz timebase I/O
- ☑ Embedded web server & iOS app
- ☑ USB flash data storage port
- ☑ HDMI video output
- ☑ GPIB, RS-232, Ethernet and USB communication

It's everything you could want in a lock-in — and then some!

DC to 4 MHz (SR865A)
DC to 500 kHz (SR860)
2.5 nV/√Hz input noise
Fast time constants

SR865A 4 MHz lock-in ... $7950 (U.S. list)
SR860 500 kHz lock-in ... $6495 (U.S. list)

SRS *Stanford Research Systems*

www.thinksrs.com/products/lockin.htm
Tel: (408)744-9040

YOUR INSIDE LOOK AT THE LATEST DISCOVERIES IN ASTRONOMY.

The virtual 237th meeting of the American Astronomical Society, held jointly with the AAS Historical Astronomy Division and High Energy Astrophysics Division, will feature a robust science program of prize and invited talks by distinguished astronomers, daily press conferences, a virtual exhibit hall, and a wide variety of short talks, digital interactive iPosters, and iPoster-Plus presentations combining talks and iPosters.

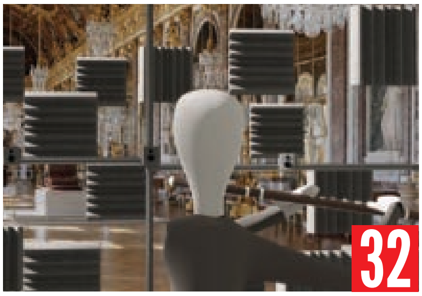
#AAS237 #globalastronomy

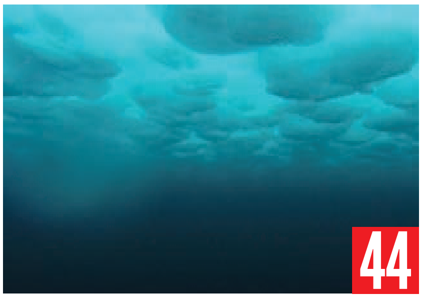
AAS | 237

237TH MEETING OF THE AMERICAN ASTRONOMICAL SOCIETY
VIRTUALLY ANYWHERE **11–15 JANUARY 2021**

REGISTER TODAY
aas.org/meetings/aas237

Pluto image courtesy of NASA, Johns Hopkins University Applied Physics Laboratory & Southwest Research Institute.





PHYSICS TODAY

December 2020 | volume 73 number 12

32 Exploring cultural heritage through acoustic digital reconstructions

Brian F. G. Katz, Damian Murphy, and Angelo Farina

Simulating the acoustics of destroyed or altered amphitheaters, cathedrals, and other architectural sites re-creates their sonic grandeur.

38 Ultrasound in air

Timothy G. Leighton

Experimental studies of the underlying physics are difficult when the only sensors reporting contemporaneous data are human beings.

44 Ocean acoustics in the changing Arctic

Peter F. Worcester and Megan S. Ballard

Recent changes in ice cover and ocean stratification have been so large that acoustic measurements made during the Cold War no longer reflect current conditions.

50 Computational phonogram archiving

Michael Blaß, Jost Leonhardt Fischer, and Niko Plath

A general approach to analyzing audio files makes it possible to re-create the sound of ancient instruments, identify cross-cultural musical properties, and more.

ON THE COVER: Experiences are often shaped by the sounds that accompany them. The mural shown here captures the joyful spirit of a jazz band leading a march for peace in East London in 1983. Although the soundtrack of that day now exists only in memory, acousticians are learning to preserve other aural experiences. On **page 32**, Brian Katz, Damian Murphy, and Angelo Farina discuss acoustic reconstructions of historical sites. And on **page 50**, Michael Blaß, Jost Leonhardt Fischer, and Niko Plath describe the re-creation of sounds from ancient instruments. (Detail from *Hackney Peace Carnival*, 1985, Ray Walker; image from marc zakian/Alamy Stock Photo.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org

TAN LIU/IONQ

Gap year

After graduating with a bachelor's degree in physics last year, Chunyang Ding chose to pursue a job in industry rather than start graduate school right away. Ding argues in a commentary that although gap years aren't for everyone, they can provide valuable experiences for aspiring physical scientists.

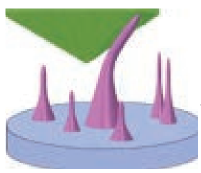
physicstoday.org/Dec2020a

TRIUMF

New APS CEO

In January theoretical physicist Jonathan Bagger will become the CEO of the American Physical Society. In an interview he talks about remote meetings, the challenges and opportunities facing the organization and the physics community, and his priorities as he begins his new role.

physicstoday.org/Dec2020b



Z. SHI ET AL., PNAS 2020

Metallic diamond

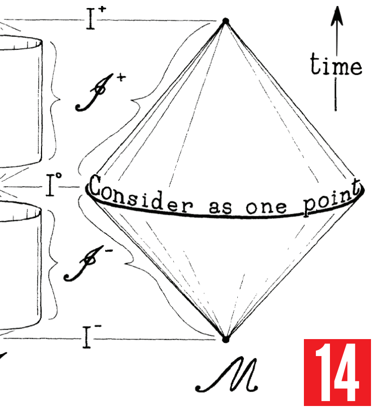
Etching diamond into fine, 300-nm-wide needles transforms the normally inelastic, brittle material into a bendable form that can take on enormous stress. Now a study demonstrates that elastic strain can alter the bandgap of the diamond nanoneedles from that of an insulator to that of a metal.

physicstoday.org/Dec2020c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 73, number 12. 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.

Copyright © 2020, American Institute of Physics. Single copies of individual articles may be made for private use or research. Authorization is given to copy articles beyond the free use permitted under US Copyright Law, provided that the copying fee of $30.00 per copy per article is paid to the Copyright Clearance Center, 222 Rosewood Dr, Danvers, MA 01923. For articles published before 1978, the copying fee is $0.25 per article. Authorization does not extend to systematic or multiple reproduction or to republication in any form. In all such cases, specific written permission from AIP must be obtained. Send requests for permission to AIP Office of Rights and Permissions, 1305 Walt Whitman Rd, Suite 300, Melville, NY 11747-4300; phone +1 516 576-2268; email rights@aip.org.

PHYSICS TODAY
www.physicstoday.org



DEPARTMENTS

8 From the editor

10 Readers' forum

Letters

14 Search & discovery

The theory of black hole formation shares the Nobel Prize in Physics • Nobel Prize in Physics honors the discovery of a supermassive compact object at the heart of the Milky Way • Charpentier, Doudna win chemistry Nobel for development of CRISPR-Cas genome editing

24 Issues & events

High radiation dose rates may improve cancer therapy • Hydrogen-powered aircraft may be getting a lift

56 Books

The importance of hypothesis — *Bernard Lightman* • Looking back on a modern scientific revolution — *Francis-Yan Cyr-Racine* • New books & media

60 New products

Focus on software, data acquisition, and instrumentation

64 Obituaries

Harry Suhl

66 Quick study

Making robots microscopic — *Marc Miskin*

68 Back scatter

Portrait of a physicist

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

Board of Directors: David J. Helfand (Chair), Michael H. Moloney (CEO), Judy R. Dubno (Corporate Secretary), Susan K. Avery, Susan Burkett, Bruce H. Curran, Eric M. Furst, Jack G. Hehn, John Kent (Treasurer), Allison Macfarlane, Michael Morgan, Elizabeth Nolan, Tyrone M. Porter, Efrain E. Rodriguez, Nathan Sanders, James W. Taylor, Charles E. Woodward.

Officers: Michael H. Moloney (CEO), Steve Mackwell (DEO), Gigi Swartz (CFAO).

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

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

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

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

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

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

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

Assistant editor
Cynthia B. Cummings

Editorial assistant
Tonya Gary

Contributing editors
Rachel Berkowitz
K. Jae
Andreas Mandelis

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

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

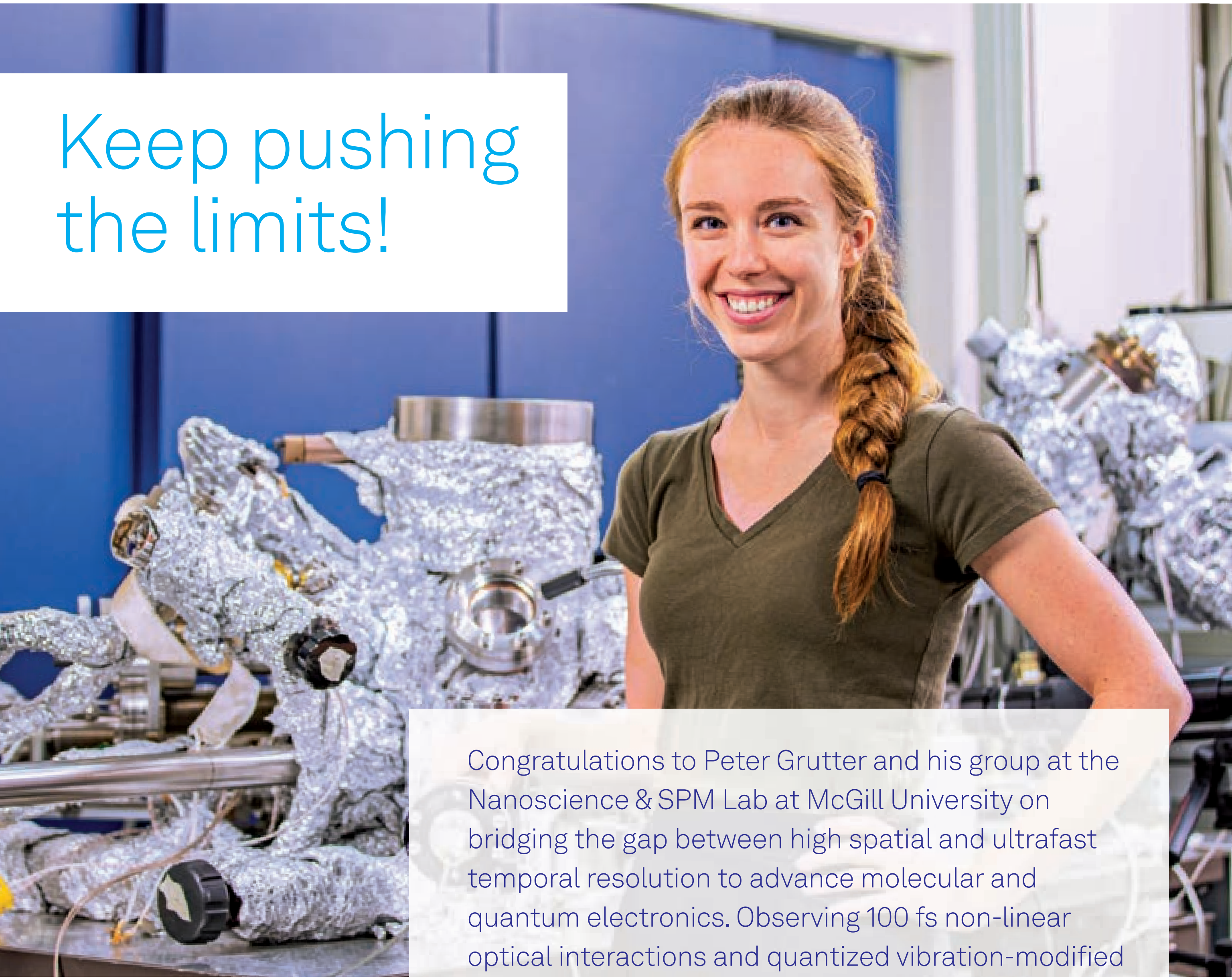
PhysicsToday **@physicstoday**

AIP | American Institute of Physics

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

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

Keep pushing the limits!



Megan Cowie, Nanoscience & SPM Group, McGill University

Congratulations to Peter Grutter and his group at the Nanoscience & SPM Lab at McGill University on bridging the gap between high spatial and ultrafast temporal resolution to advance molecular and quantum electronics. Observing 100 fs non-linear optical interactions and quantized vibration-modified electron transfer in single molecules with AFM are impressive achievements that set new standards at the forefront of scientific research.

We are excited to continue our collaboration and look forward to finding new ways of using lock-in amplifiers and boxcar averagers to push the limits of SPM applications.

Zurich Instruments

www.zhinst.com

Your Application. Measured.

FROM THE EDITOR

Sound affects

Charles Day

The variety of waves is astonishing. Even if you count only the ones named after people, the list is long: Alfvén, Bloch, Dyakonov, Faraday, Gerstner, Kelvin, Lamb, Langmuir, Love, Mach, Rayleigh, Rossby, Stokes, and Tollmien–Schlichting. The swaying motion of the Rayleigh wave, the snaking motion of the Love wave, the planetary scale of the Rossby wave—to pick just three—are beguiling.

By contrast, when sound propagates through a gas or a liquid, it does so by way of the humblest of waves, the longitudinal. But in what is perhaps the greatest irony in physics, those longitudinal waves underlie one of the richest and most varied of physical sciences: acoustics. If you browse a recent issue of the *Journal of the Acoustical Society of America*, you'll see what I mean. The table of contents typically lists around 10 different sections, which include acoustical oceanography, animal bioacoustics, biomedical acoustics, noise, and speech communication—to pick just five.

In celebration of the richness and importance of acoustics, this issue of PHYSICS TODAY is devoted to the International Year of Sound, which kicked off on 31 January with a ceremony at Sorbonne University's Grand Amphithéâtre in Paris. The four feature articles in this issue were chosen with the help of the current president of the Acoustical Society of America, Diane Kewley-Port of Indiana University.

The richness of acoustics is manifest in the news stories I've written for this magazine. That is revealing, because when contributors to the Search and Discovery department choose papers to cover, they prioritize significance of research followed by variety of topic in the issue. We don't set out to build personal bibliographies of wide variety.

My first acoustical Search story (June 2001) reported on research that identified the neurological basis of sound location in barn owls. Next came overcoming the diffraction limit with time-reversed ultrasound (December 2002), mediating the delivery of therapeutic genes to heart muscle with targeted ultrasound (December 2005), mitigating hospital noise with acoustic paneling (August 2007), and the confirmation, using *ex vivo* and virtual cells, of a theory of cochlear amplification (June 2010).

I stopped being a regular contributor to Search when I became online editor in June 2010. But I continue to write news stories for the magazine's website. The most recent was about an ingenious experiment that determined which of two aural cues—interaural level differences and interaural time differences—crocodiles use to locate squealing hatchlings, approaching prey, and other sources of sound. One of the experiment's subjects appears in the photo.

Although I never learned to play a musical instrument with any skill, music has remained for me one of life's dearest pleasures, the more so during the current pandemic. Listening to music at home is a safe and stimulating pastime. I favor CDs. Among my recent purchases are Anton Bruckner's string quartet and quintet; James Brown's 1973 swansong, *The Payback*; jazz musician Zbigniew Namysłowski's 1973 album, *Winobranie*; and the Cocteau Twins' 1984 album, *Treasure*.

I list my CDs not to show off my eclectic taste, but to make a point. I've known about and delighted in the variety of musical forms ever since I was a nipper in North Wales. But it took reporting on the physical sciences for me to appreciate the variety of acoustical research, by which time I was well into my 30s. As the organizers of the International Year of Sound put it, "sound is omnipresent in our lives." Understanding of sound and its applications, however, is far from omnipresent. The more IYS and other initiatives engage the public in the science of sound, the better. PT

DO YOU SUPPORT SCIENCE?

This holiday season, as you plan your charitable giving, we invite you to consider the American Institute of Physics, and its signature programs: The Center for History of Physics, The Niels Bohr Library and Archives and support of students fostered through Sigma Pi Sigma and the Society of Physics Students.

A gift through the AIP Foundation to these programs helps to keep the broad enterprise of science strong and vibrant today and into the future.

Visit: **foundation.aip.org** to learn more!

Planning your philanthropy in 2020

Here are some considerations for year-end giving in 2020:

- For those who do not itemize their tax returns, the CARES Act allows for a deduction of up to $300 for charitable giving
- The CARES Act allows itemizers to deduct charitable contributions up to 100% of their AGI in 2020, providing a way to support charitable priorities while also reducing tax exposure
- If you are 70½ or older, you may wish to consider a Qualified Charitable Distribution (QCD) as a way of supporting the AIP or other charities while also reducing tax exposure related to IRA distributions—the QCD is a payment from your IRA transmitted at your request directly from your IRA custodian to charity (up to a maximum annual amount of $100,000)
- Appreciated securities can provide advantages to donors who do not wish to make an outright contribution

We recommend that you consult your financial advisor to learn more about options available to you.

To learn more about ways you can have an impact by giving through the AIP Foundation, visit us today: **foundation.aip.org**

Questions? Email us: **aipfoundation@aip.org**

AIP FOUNDATION

Early solar eclipse images and tests of relativity

In a letter in the June 2020 issue of PHYSICS TODAY (page 12), Robert McAdory posed the question of whether photographic plates taken at earlier solar eclipses might have been of use to the famous British eclipse expeditions of 1919. Those expeditions tested Albert Einstein's prediction that the positions of stars seen near the Sun are affected by the gravitational deflection of light. Deborah Kent's article about the solar eclipse of 1869 (PHYSICS TODAY, August 2019, page 46) discussed very early photographic plates of the Sun. Because we have both written about astronomers' testing of Einstein's theory of general relativity, we wanted to expand on that interesting discussion.

Stars almost certainly did not appear on plates taken in 1869, because the wet-plate photography in use then was not very sensitive. Generally, only bright objects like the Sun or Moon appeared on photographs of the sky. Not until the development and popularization of dry-plate photography in the 1880s did stellar photography become possible.

Some star images were obtained during eclipses before 1919, and efforts were made to use those plates to retroactively test Einstein's light-deflection prediction. However, none of the images were clear enough for measuring the small light-deflection effect, essentially because of the following confounding factors.

THE GREAT PROMINENCE VISIBLE HERE is from the famous 1919 eclipse. The photo was taken by Arthur Stanley Eddington using an Astrographic lens on the island of Principe. This image was worthless for testing Einstein's light deflection effect. Clouds occluded the stars near the Sun but prevented overexposure of the bright prominence. (Courtesy of Charlie Johnson.)

▶ **Too few stars were imaged.** Frank Dyson and Charles Davidson, the two astronomers mainly responsible for data analysis of the plates taken in Sobral, Brazil, in 1919, had examined plates from a 1905 eclipse taken with one of the lenses that would be used in 1919. They had found only two stellar images. But they anticipated that the same setup would work in 1919, since the star field was much richer in bright stars. For subsequent eclipses, astronomers took very long exposures in order to image enough stars. The 1922 eclipse exposures, for example, were about one minute long.

▶ **The field of view was too narrow.** Most eclipse plates are framed tightly on the Sun and its corona and do not include other stars at all. Plates taken in search of the nonexistent planet Vulcan were different, but they tended to place the Sun in the corner of a plate, when its being in the center is optimum for the light-deflection experiment.

▶ **Tracking was inappropriate.** Because the Sun is typically the center of attention during eclipses, astronomers naturally tracked with their telescopes to keep the Sun fixed on the plate. Star images would not be so fixed, so any that appeared would sustain some streaking, which would compromise the measurement of position shifts.

▶ **No comparison plates were available.** Before 1919 there was no reason to make comparison plates of the same star field with the same equipment months before or after an eclipse. But without them, any small shift of position would be difficult to notice, let alone measure. In the early 20th century, positions for the vast majority of stars were not accurately known. In fact, Dyson and Davidson were beavering away on the first all-sky photographic survey, the Carte du Ciel project. Even their routine measurements were not necessarily accurate enough to notice Einstein's predicted effect.

American astronomer Heber Curtis (who attempted the Einstein experiment during a 1918 eclipse) did try to analyze plates from the 1900 eclipse, using plates taken in 1919 for comparison. He measured the positions of six stars visible on the plates, but with the 19-year time lapse, he could not rule out proper motion (the projected motion of the stars in the Milky Way) as an explanation of any shift in position. That would be especially true of the particular star field, since the Hyades stars—in which the Sun is located on 28 and 29 May, the dates of the 1900 and 1919 eclipses—are close to our solar system and exhibit large proper motions.

For his six stars, Curtis did use rectilinear coordinates from the Paris zone of the Carte du Ciel project, and he believed they supported his contention that the predicted light-deflection effect was not real. But his data were of poor quality, and William Wallace Campbell, his collaborator and employer at the Lick Observatory, declined to publish the results.

In his letter, McAdory asked whether any astronomers, independent of Einstein, suspected the existence of a shift of star positions near the Sun. One who did believe was Leopold Courvoisier, a Swiss astronomer at the Babelsberg Observatory in Berlin and a colleague of Einstein collaborator Erwin Finlay Freundlich. However, Courvoisier thought the effect extended much farther from the Sun and could be observed without the need for eclipse expeditions. The shift was essentially a seasonal one, he believed, bigger than the one Einstein called for, with the Sun at its center. He was a staunch antirelativist who hated Einstein's theory, and he attributed the effect to the solar system's motion through the ether.

Daniel Kennefick
(danielk@uark.edu)
University of Arkansas
Fayetteville

Jeffrey Crelinsten
(jcrelinsten@impactg.com)
The Impact Group
Toronto

CONTACT PHYSICS TODAY

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

Eliminating the GRE

A musician would not take a multiple-choice test to join a band. Yet such tests are how many physics and astronomy graduate programs select students. As a result, departments unintentionally limit the talent that enters the field.

With testing centers closed because of the coronavirus and with concerns about the fairness or the lack[1] of online alternatives, physics and astronomy graduate programs have temporarily stopped using entrance exams to make admissions decisions. I suggest that become a permanent change.

Many faculty members and students believe that the physics Graduate Record Examination (GRE) is one of the most important aspects of graduate admissions.[2] Many programs require a certain score for admission. In theory, a standardized test provides an objective measure for comparing applicants. Everyone takes the same test, so it allows candidates the ability to distinguish themselves by scoring well, or so the argument goes.

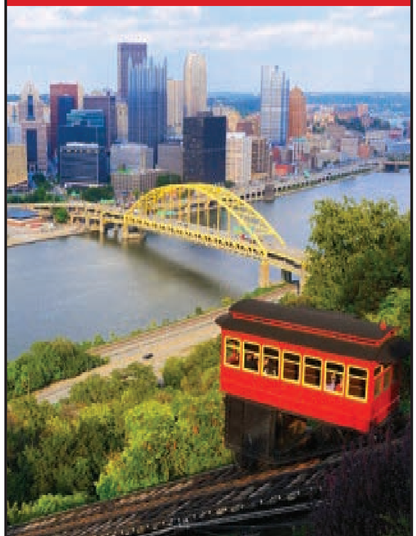
But that's not exactly right. With Marcos Caballero, I conducted a study[3] of more than 2500 applicants to five physics programs; we found that doing well on the physics GRE did not result in a higher chance of admission than earning high grades did. In fact, nine times as many applicants are likely to be hurt by a poor physics GRE score despite high grades as are likely to benefit from a high score despite low grades. That discrepancy is in addition to the known test disparities based on gender and race: Women and people of color score lower on the GREs than their white male peers. The differences in scores are because of circumstances, not ability.

Take, for example, stereotype threat. Imagine a Black student hearing her peers say people like her are not good at physics or don't belong in it. If she does poorly on the test, her peers might take it as evidence that they were correct. Now, besides worrying about the test, she worries about proving herself. As a result, she cannot concentrate on the test, and she scores worse than students who didn't worry about stereotypes. The test is hardly a fair method of comparing applicants.

Then there is the expense. While it affects all students, it particularly hurts

M&M 2021
MICROSCOPY & MICROANALYSIS

August 1-5
Pittsburgh, PA



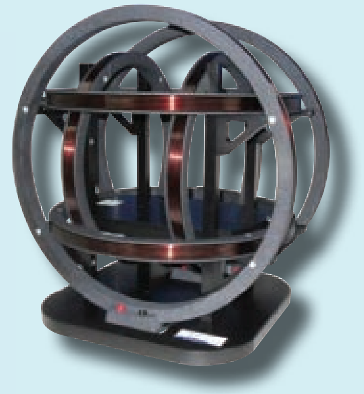
2021 Call for Submissions

Deadline:
February 18, 2021

Submissions portal opens:
December 7, 2020

Go to
www.microscopy.org/MandM/2021
for portal access and up-to-date meeting information

MSA Microscopy Society of America
MAS Microanalysis Society


Helmholtz Coil Systems
- 350mm to 2m diameter coils
- Orthogonality correction using PA1
- Active compensation using CU2
- Control software available


Mag-13 Three-axis Magnetometers
- Noise levels down to <4pTrms/√Hz at 1Hz
- Measuring ranges from ±60 to ±1000μT
- Bandwidth of DC to 3kHz

US distributor:
GMW Associates
Telephone: +1 (650) 802-8292
www.gmw.com

Bartington® Instruments
www.bartington.com

those who come from low-income households. Taking the physics GRE requires a $150 fee. If applicants apply to more than four programs, they pay an extra $27 per school to share the test results. Travel costs must be considered as well, because some students may not be near a GRE testing center. Arizona and Nevada, for example, each have only one such site in the state.

If the test provided useful information for admission, perhaps those problems would be tolerable. Medicines have side effects, but medical professionals prescribe them when the benefits outweigh the risks. Such is not the case with the physics GRE. A study of nearly 4000 physics PhD students[4] found only a minor difference between the fractions of high- and low-scoring applicants who completed their PhDs.

Earning a PhD is only one part of being a successful graduate student. Research success is just as important, but the GRE is not helpful in that respect either. One study looked at 149 past recipients of the most prestigious astronomy research fellowships and found that many of them had scored below a typical admissions cut-off on the physics GRE.[5]

Departments would have already removed the physics GRE requirement if doing so were simple. Some faculty members advocate keeping the GRE because they believe it measures something useful. Perhaps departments could compromise and allow applicants to choose whether to submit scores. It seems like having the choice would benefit those who score poorly or who cannot afford to take the test. However, applicants have varying ideas about what "test-optional" means.

As part of a larger 2020 study on physics GRE requirements creating an uneven playing field, researchers from Rochester Institute of Technology interviewed 19 graduate students who had applied to test-optional programs. The 10 women reported submitting their scores, regardless of how well they did, while the 9 men had submitted scores only if they were stellar. So by becoming test-optional, the admissions process may become further biased.

To make the physics graduate school admissions process more equitable, departments must expunge the physics GRE and reenvision how they admit applicants.

Departments could start by admitting students based on their potential instead of their achievement, as recommended in the *Final Report of the 2018 AAS Task Force on Diversity and Inclusion in Astronomy Graduate Education*. Comparing achievements is of limited use when not everyone has the same opportunities.

When conducting research, scientists consider confounding factors that could invalidate their experiments. Admissions committees must do the same when admitting applicants to physics and astronomy programs. Including in evaluations academic achievement and noncognitive skills such as conscientiousness and perseverance can help departments select applicants who can be successful in their programs and do so equitably.

Furthermore, according to an Ohio State University survey, programs that have removed the physics GRE have had more applicants, especially among those identifying as Black, Latinx, and Native American.

The coronavirus has caused departments to rethink how to evaluate applicants. What a great opportunity to remove one of the least equitable parts of the process!

References

1. J. Hu, *Science* (2020), doi:10.1126/science.caredit.abd4989; ETS, "GRE Subject Test Administrations Canceled for September 12 and October 17" (20 August 2020).
2. G. Potvin, D. Chari, T. Hodapp, *Phys. Rev. Phys. Educ. Res.* **13**, 020142 (2017); D. Chari, G. Potvin, *Phys. Rev. Phys. Educ. Res.* **15**, 023101 (2019).
3. N. T. Young, M. D. Caballero, https://arXiv.org/abs/2008.10712.
4. C. W. Miller et al., *Sci. Adv.* **5**, eaat7550 (2019).
5. E. M. Levesque, R. Bezanson, G. R. Tremblay, http://arxiv.org/abs/1512.03709.

Nick Young
(youngn18@msu.edu)
Michigan State University
East Lansing

Corrections

October 2020, page 40—In the caption for figure 1, the increase in the number of physics PhDs at US universities over the past 15 years should be 75%.

November 2020, page 28—Daniel Bernoulli, not David, devised a model to predict the benefits of smallpox inoculations. In box 1, the negative binomial distribution should be $P(x; R_0, k) = [\Gamma(k + x)/\Gamma(k + 1)\Gamma(x)]p^k(1 - p)^x$. PT

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Associate/Full Professor Position
in Neutron Scattering for Quantum Materials

The Department of Physics and Astronomy, in collaboration with the Department of Materials Science and Engineering at the University of Tennessee, Knoxville, invites applications for a faculty position at the Associate/Full Professor level, beginning August 1, 2021. The successful candidate will hold a joint appointment with both departments. This search aims to strengthen our efforts in the understanding, characterization, and development of new quantum materials, which requires tight collaboration between the two departments. Additional hires in quantum materials and/or quantum information are anticipated.

Both departments hold a deep commitment toward developing and promoting a culturally inclusive community. We strongly encourage applications from members of groups that are underrepresented in science and technology.

Candidates should have a PhD in physics or closely related field and an exemplary record of research in experimental neutron scattering for quantum materials. They should have experience in the modelling and interpretation of neutron data, interface effectively with computational physicists, and demonstrate a keen ability to communicate with theorists. The successful applicant will establish an externally-funded research program, provide interdisciplinary training for graduate students and postdocs, and contribute to the teaching mission of both departments. We expect the candidate to participate in initiatives and coordinated research efforts within the framework of the newly-developed cluster in 'Quantum Materials for Future Technologies.'

Applicants should upload a CV, list of publications, description of research and teaching experience, proposed research program, and contact information of three references at **https://apply.interfolio.com/80082**. Only electronic applications will be considered. Acceptable file formats are .pdf or .docx. We will start reviewing applications by January 15, 2021.

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.

THE UNIVERSITY OF TENNESSEE KNOXVILLE

Assistant Professor Position
in Theoretical Condensed Matter Physics

The Department of Physics and Astronomy, in collaboration with the Department of Materials Science and Engineering at the University of Tennessee, Knoxville, invites applications for a tenure-track faculty position, beginning August 1, 2021. The successful candidate will have a primary appointment in the Department of Physics and Astronomy with potentially a joint appointment in the Department of Materials Science and Engineering. This search aims to strengthen our efforts in the understanding, characterization, and development of new quantum materials, which requires tight collaboration between the two departments. Additional hires in the areas of quantum materials and/or quantum information are anticipated.

Both departments hold a deep commitment toward developing and promoting a culturally inclusive community. We strongly encourage applications from members of groups that are underrepresented in science and technology.

Candidates should have a PhD in physics or a closely related field and an exceptional research record in the areas of theoretical topological quantum materials including topological superconductors, dynamics and manipulation of quantum systems, solid state implementation of quantum information, and quantum device physics. The successful candidate will establish an externally-funded research program, provide interdisciplinary training for graduate students and postdocs, and contribute to the teaching mission of both departments.

Applicants should upload their CV, list of publications, description of research and teaching experience, and proposed research program at **https://apply.interfolio.com/80082**. Three letters of reference should be uploaded separately. Only electronic applications will be considered. Acceptable file formats are .pdf or .docx. We will start reviewing applications by January 15, 2021.

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.

The theory of black hole formation shares the Nobel Prize in Physics

Roger Penrose proved that black holes could and should form in the universe. The topological picture he used in that proof has become foundational to general relativity.

FESTIVAL DELLA SCIENZA/CC BY-SA 2.0

In a 1969 paper, Roger Penrose wrote, "I only wish to make a plea for black holes to be taken seriously and their consequences to be explored in full detail. For who is to say, without careful study, that they cannot play some important part in the shaping of observed phenomena?"[1] His success at getting researchers to consider black holes in earnest was acknowledged with half of this year's Nobel Prize in Physics.

With decades of evidence, it may be hard to imagine that physicists ever doubted their existence. That evidence includes work by Reinhard Genzel and Andrea Ghez, who received the other half of this year's prize (see the story on page 17), and the Event Horizon Telescope image released last year (see the Quick Study by Dimitrios Psaltis and Feryal Özel, PHYSICS TODAY, April 2018, page 70). General relativity, which admits the possibility of singularities, has cemented its place as the standard theory for gravitational phenomena and the basis for many cosmological models. But that fate was far from certain in the 1960s when Penrose started his work on black holes.

In a three-page paper with few equations, Penrose showed in 1965 that black holes were not a mathematical oddity but a real-life inevitability,[2] and he introduced the mathematical framework of topology to general relativity for the first time. He is "unmatched by anyone since Einstein in his contribution to our understanding of gravity," says astrophysicist Martin Rees.

The general idea

As early as the late 1700s, John Michell and Pierre-Simon Laplace had hypothesized about an object with an escape velocity faster than the speed of light, such that even particles of light couldn't leave its gravitational field. Using Newtonian mechanics, Michell and Laplace each found that a spherical mass m confined to a radius less than $2Gm/c^2$ would produce that behavior.

That same radius would appear again more than 100 years later. In 1915 Albert Einstein published his theory of general relativity, which relied on unfamiliar and difficult mathematics that he could solve only approximately. Early in 1916, while serving in the German army in World War I, Karl Schwarzschild produced the first solution to Einstein's field equations. A mathematical oddity appeared in his solution for the curved spacetime around spherically symmetric matter of mass m and radius r: Some terms vanish or diverge for $r = 0$ and $r = 2Gm/c^2$.

The so-called singularity at $r = 0$ was a point of infinite density with gravitational pull so strong that not even light could escape. But researchers weren't sure how to understand the result. Were the singularities real or an artifact of the choice of coordinates?

Twenty years later, J. Robert Oppenheimer tried to make sense of gravitational singularities by analyzing the collapse of a spherical cloud of matter. He and his student Hartland Snyder were the first to realize the meaning of the second notable r value, now known as the event horizon: The collapsing star wouldn't be able to communicate with observers beyond the horizon. But Oppenheimer's assumption of spherical symmetry aroused skepticism about whether such a circumstance would occur in real life. Would asymmetry prevent the collapse toward a single point? Could matter bounce back out? Physicists Evgeny Lifshitz, Isaak Khalatnikov, and John Wheeler each stated that such collapse wouldn't occur in realistic conditions.

Beyond that work, in the mid 1920s to mid 1950s, general relativity broadly was considered little more than a small correction to the Newtonian picture of gravitational interactions.[3] What's more, the accuracy of general relativity compared with other proposed gravitational theories was still a matter of debate.

One prediction of Einstein's theory, the bending of light by gravity, was confirmed by Arthur Eddington's astronomical observations during the solar eclipse of 1919 (see the article by Daniel Kennefick, PHYSICS TODAY, March 2009, page 37) and in another measurement in 1922 by astronomers from the Lick Observatory.

Further experimental evidence for general relativity awaited advances in technology, and other unified field theories could explain gravitational light bending. Theo-

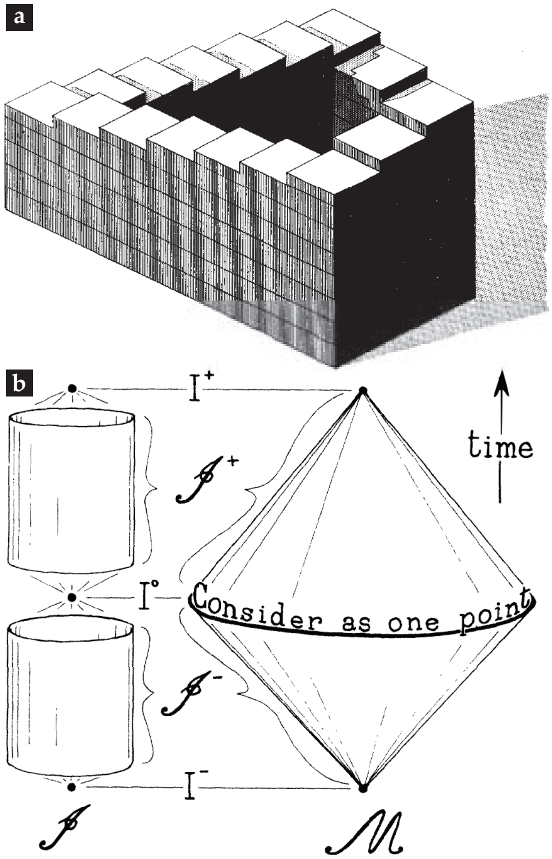


FIGURE 1. ROGER PENROSE PLAYS WITH DIMENSIONS in his impossible stairs and spacetime diagrams. **(a)** Penrose stairs are one of the impossible objects Penrose designed with his father. Although such an object cannot exist as it appears in the image, with stairs ascending in a loop, real objects can create the illusion with a carefully selected shape viewed from the right angle. (Adapted from ref. 5.) **(b)** Penrose diagrams also use a trick perspective to draw the entire universe. Through a technique known as a conformal transformation, which preserves angles but not distances, the infinite past I^-, spatial infinities I^0, infinite future I^+, and light's paths to infinity ($\mathscr{I}^+$ and $\mathscr{I}^-$) become finite in conformal spacetime $\mathcal{M}$ and can be graphed. (Adapted from ref. 6.)

retical physicists instead turned their attention to quantum mechanics, which offered clear connections to experiments and plentiful potential applications.

Physics in the other renaissance

After World War II, what general relativity theorist Clifford Will called a "renaissance of general relativity" bloomed.[4] Physics had played a pivotal role in the war and in the Cold War that followed, and the result was that it enjoyed a boost in funding and an attraction of talent. The subsequent increase in postdoc positions helped transmit ideas in the previously small and geographically dispersed field of general relativity. The field eventually became well-defined, with regular conferences and publications devoted exclusively to it.

Recent high-precision instrumentation, such as lasers and atomic clocks, enabled tests of general relativity that hadn't been feasible before. "As of the early 1960s the evidence supporting general relativity was thin at best," says Will. "There was some, but it was pretty feeble." With the ability to test aspects of gravitational theories, researchers regained interest.

After the discovery of quasars in the early 1960s, astronomers started to realize that general relativity might explain more than minor deviations from Newtonian theory. Quasars are apparently compact sources whose emissions change even on the scales of days or hours and are brighter than an entire galaxy. Their prodigious energy output couldn't be accounted for in Newtonian gravity. Around the same time, Robert Pound and Glen Rebka's gamma-ray experiments confirmed another prediction from general relativity: gravitational redshift.

On the theoretical front, mathematician Roy Kerr generalized Schwarzschild's field-equation solution to a rotating body. And Martin Kruskal and David Finkelstein's reexamination of the Schwarzschild result using different coordinates explained away the apparent singularity for a local observer at $r = 2Gm/c^2$. Finkelstein's lectures on the Schwarzschild solution are part of what inspired Penrose to pick up the problem in 1964.

Roger that

"Unlike most relativists, especially those in the US, Penrose came from a background in mathematics rather than physics," says Rees. "That is why he had such distinctive influence in the 1960s."

Penrose earned a degree in mathematics from University College London, where his father was a professor of human genetics, and went on to receive a PhD at the University of Cambridge in 1957. As a student Penrose was geometrical in his thinking, even compared to his peers in mathematics; for example, he formulated pictures as an alternative way to write complicated mathematical expressions, such as tensors. At Cambridge he attended lectures outside his field, notably those on relativity by Hermann Bondi and quantum mechanics by Paul Dirac.

During Penrose's university years, his interest in astronomy increased due to Fred Hoyle's BBC radio talks that covered astronomy and cosmology, including steady-state theory, which asserts that the universe is the same at any time and place. (Getting the theory to work in an expanding universe requires the constant production of new matter; see the article by Geoffrey Burbidge, Fred Hoyle, and Jayant V. Narlikar, PHYSICS TODAY, April 1999, page 38.) Hoyle mentioned that a galaxy traveling the speed of light would disappear from the visible universe. When Penrose considered the problem in terms of light cones, he couldn't make sense of it. That puzzle was one of the first things he discussed when he met cosmologist Dennis Sciama, who encouraged Penrose's interest in physics.

As a graduate student, Penrose also stumbled on the work of artist M. C. Escher. He was inspired to try his hand at similar, mathematically inspired designs, and he and his father devised their own impossible objects, some of which were published in the *British Journal of Psychology*.[5] One example, known as Penrose stairs, is shown in figure 1a. A later outgrowth of that interest was his design of an aperiodic tile pattern called Penrose tilings, the first nonrepeating structure to feature two shapes. The pattern later appeared in quasicrystals, the subject of the 2011 Nobel Prize in Chemistry (see PHYSICS TODAY, December 2011, page 17).

"It's clear that he not only thinks very deeply, but in a very unusual and specially geometric way," says Rees. "His notebooks look like those of Leonardo da Vinci!"

Finite infinities

As advances in experimental technology opened new views of the universe, so did theoretical tools. Among them, Penrose diagrams, published in 1962, were particularly powerful.[6] The diagrams are a visual representation of all spacetime, and after Sciama's graduate student Brandon Carter helped popularize the diagrams, they became ubiquitous in black hole physics of

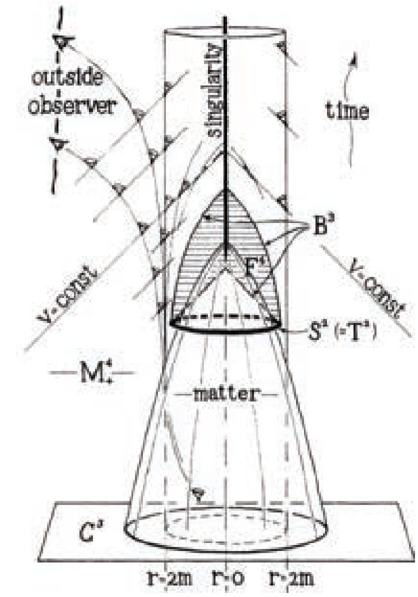


FIGURE 2. MATTER COLLAPSES INTO A BLACK HOLE. As the radius *r* of matter with mass *m* dips below 2*m*, a distance known as the event horizon (cylinder), a trapped surface (black ring at center) emerges in the empty space around the matter. On that closed surface, even light can't move away from the singularity, and the surface thus signals the inevitable formation of a black hole. All the while an outside observer never sees matter cross the event horizon. Although this sketch is for a spherically symmetric mass, Penrose proved the result would be true for any shape. (Adapted from ref. 2.)

the 1960s. They remain a standard in general relativity education and research.

The diagrams are derived from Hermann Minkowski's 1907 depictions of spacetime with time as the vertical axis, one spatial dimension as the horizontal axis, and light traveling at 45°. Because an observer can't travel faster than light, the spacetime available to them is defined by a 45° cone. But the meaning and interpretations of those diagrams weren't well understood or agreed upon.

Physicists often consider the behavior of fields at infinity, but a point or observer at infinity doesn't exist. Penrose found a way to concretize infinity through conformal transformations—that is, angle-preserving ratios of distances—of Minkowski space. After that transformation, infinity becomes the boundary of a finite region. "You can literally draw a universe," says historian of science Aaron Wright, "and that was just not physically possible before that new technique."

The infinite past I^-, spatial infinities I^0, and infinite future I^+ transform into the point at the bottom, the ring at the center, and the point at the top right in figure 1b for conformal spacetime $\mathcal{M}$. The bounds of the cone show light's path to infinity, traveling along the 45° lines.

Penrose diagrams are similar in spirit to the puzzles and impossible objects Penrose designed with his father. As your eye moves around the diagram, the spatial relationships between you and the page change. For the diagram, the distance scales near the infinite edges appear dramatically smaller than at the center.

The drawings had "the same impact in our field as Feynman diagrams had in particle physics," says Will. They didn't just visualize what's happening. The drawings are formal mathematical objects that can be used to perform calculations and obtain quantitative results without pages of equations.

It's a trap

Diagrammatic analysis would serve Penrose well. In 1964 John Wheeler started reconsidering gravitational collapse and discussed the matter with him. Penrose approached the problem without assuming spherical symmetry or other idealized assumptions, as had been done previously.

Penrose's key insights, published in 1965, are summarized in figure 2. The figure shows two spatial dimensions with time as the third axis. Starting from the bottom, matter collapses to $r = 0$, and from the matter's perspective, it passes over the horizon in finite time and doesn't experience anything special. But to an outside observer the matter will be eternally collapsing toward the event horizon (shown as a cylinder at $r = 2m$, given in units for which $G = c = 1$) but never reaching or crossing it.

Penrose identified a point of no return—the formation of what he called a trapped surface—after which the collapse into a singularity, or black hole, is inevitable and unavoidable. From the collapsing matter's perspective, once it is past the event horizon, it is surrounded by empty space, which hosts the trapped surface. In that bubble of spacetime, space and time dimensions swap properties: Space becomes timelike, and trajectories can move in only one direction. The existence of a trapped surface thus leads inevitably to a singularity.

A trapped surface forms even if the matter distribution lacks a sharp boundary, is asymmetric, or rotates. In fact, Penrose's discussion of star collapse relies on few assumptions included in Einstein's field equations; using a leading competitor gravitational theory of the time, posited by Carl Brans and Robert Dicke based on work by Pascual Jordan, wouldn't qualitatively change his argument. Once enough mass occupies a given volume, pressure can't counteract the gravity, and a black hole will form.

Causal effects

After Penrose's analysis of gravitational collapse, black holes became the prevailing explanation for quasars. His research helped turn scientific opinion on black holes from unlikely theoretical entities to a plausible explanation of quasars, blazars, and other active galactic nuclei (for more, see the article by Neil Gehrels and Jacques Paul, PHYSICS TODAY, February 1998, page 26). The emission from quasars, for example, is understood as the release of rest-mass energy as a result of gravitational stresses and friction as matter falls into a black hole.

Penrose proceeded to elucidate the structure, dynamics, and nature of black holes. For example, in 1971 he realized that a black hole's rotational energy could be physically extracted from the black hole. For any rotating mass, spacetime swirls around it. For the case of a rotating black hole, any observer cannot avoid rotating with that swirling spacetime. If a projectile swirling with spacetime splits in two such that one part escapes the black hole, the escaping part can have energy higher than the original projectile because of the black hole's rotational energy.

Penrose also went on to collaborate with Stephen Hawking, a student of Sciama's, and their application of similar trapped-surface reasoning to cosmological singularities supported the existence of a past singularity—the Big Bang.

"I think the Nobel Prize is long, long overdue" for Penrose and for general relativity, says theoretical physicist Lee Smolin. (Even Einstein won for the photoelectric effect, not general relativity.) Smolin emphasizes the influence Penrose has had on the mathematics used in general relativity, such as conformal structure.

With the recent high-profile results related to general relativity, such as the observation of gravitational waves (subject of the 2017 Nobel Prize in Physics; see

PHYSICS TODAY, December 2017, page 16), "I think we're on the verge of a totally new renaissance," says Will, "and this one is going to involve spreading general relativity to the strong-gravity dynamical regime." Most of the results from the post–World War II renaissance involved weak fields and slow motions relative to the speed of light. Strong fields emerge in, say, two black holes circling one another at half the speed of light with their horizons touching. How will matter behave then?

Heather M. Hill

References

1. R. Penrose, *Riv. Nuovo Cimento* **1**, 252 (1969).
2. R. Penrose, *Phys. Rev. Lett.* **14**, 57 (1965).
3. A. Blum, D. Giulini, R. Lalli, J. Renn, *Eur. Phys. J. H* **42**, 95 (2017).
4. C. M. Will, *Was Einstein Right? Putting General Relativity to the Test*, Basic Books (1986).
5. L. S. Penrose, R. Penrose, *Br. J. Psychol.* **49**, 31 (1958).
6. R. Penrose, *Phys. Rev. Lett.* **10**, 66 (1963); A. Wright, *Hist. Stud. Nat. Sci.* **44**, 99 (2014).

Nobel Prize in Physics honors the discovery of a supermassive compact object at the heart of the Milky Way

By tracking the orbits of stars close to the galactic center, Reinhard Genzel and Andrea Ghez ruled out all possibilities besides a black hole.

Reinhard Genzel

MAX PLANCK INSTITUTE FOR EXTRATERRESTRIAL PHYSICS, GARCHING

Andrea Ghez

CHRISTOPHER DIBBLE

When newly graduated Karl Jansky joined Bell Laboratories in 1928, his first assignment was to determine the source of static that plagued transatlantic radio transmissions. Having built a rotating antenna, he found a radio signal emanating from the center of the Milky Way, in the direction of the constellation Sagittarius. On 5 May 1933 the *New York Times* reported Jansky's discovery of "star noise." The radio source later became known as Sagittarius A.

Although Jansky (1905–50) did not live long enough to learn the implication of his discovery, those radio waves were one of the first observational forays that later revealed the black hole sitting at the galactic center. The next decades yielded theoretical and observational results that explained radio emissions coming from distant galaxies and uncovered what lies at the heart of our own galaxy.

Reinhard Genzel, Andrea Ghez, and their collaborators provided experimental evidence that pinned down the source of the radio signals. Their painstaking measurements of the orbital motions of stars near the galactic center conformed to general relativity predictions of objects swirling around a black hole. The Nobel Prize committee awarded half of this year's physics prize to Genzel and Ghez, "for their discovery of a supermassive compact object at the center of our galaxy." Their insights set a new course for testing subtle effects of general relativity.

It started with a quasar

Radio astronomy was in its infancy in the 1950s when astronomers were puzzled by all-sky surveys that revealed radio emissions from compact, apparently extragalactic locales with no corresponding objects in the visible. Those mysterious emissions were dubbed quasi-stellar objects, or quasars. (See the article by Hong-Yee Chiu, PHYSICS TODAY, May 1964, page 21.) Supermassive black holes with masses of 10^6 to 10^9 times that of the Sun were proposed as an explanation, and one was even posited to sit at the center of our galaxy—a relic of a quasar's younger, hotter phase.

In 1971 Donald Lynden-Bell and Martin Rees applied the black hole model of quasars to predict the spectroscopic signatures of dust and stars moving around a putative supermassive black hole at the center of the Milky Way. Their predictions inspired astronomers to look in our cosmic backyard to establish the existence of supermassive black holes.

Jansky traced radio waves to the constellation of Sagittarius in 1933, but funding shortages and office politics prevented him from further investigating their origin. It wasn't until 1974, at the National Radio Astronomy Observatory, that Robert Brown and Bruce Balick pinpointed an even stronger and more compact source of radio emissions from the center of the Milky Way, within the source reported by Jansky, near the border of the constellations Sagittarius and Scorpius.[1] Brown named that compact radio source Sagittarius A*, or Sgr A*.

Measuring the velocities of stars and gas clouds orbiting Sgr A* would ultimately determine the mass and radius of the source. However, tools for doing so were in their infancy. Thick clouds of dust obscure the galactic center in visible light, so sensitive spectroscopy in the far-IR was needed to tease out both the chemical structure and the motion of gases swirling around Sgr A*.

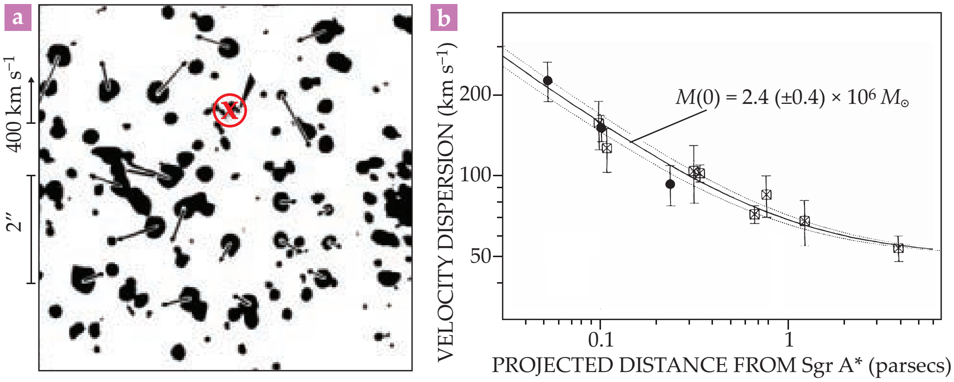


FIGURE 1. STARS MOVE NEAR THE GALACTIC CENTER. (a) Vectors indicate stellar motions in the immediate vicinity of radio source Sagittarius A* (Sgr A*), denoted by the red X, as determined by a four-year survey. **(b)** A star's velocity is a function of its distance from Sgr A*. Solid circles indicate the velocity dispersion estimated from stars' proper motions, the apparent shifts in their positions. Open rectangles indicate velocity dispersion estimated from stars' line-of-sight velocities. Curves indicate the velocity behavior expected for a single massive point source with 2.4 million $M_\odot$. (Adapted from ref. 4.)

"That's where my second father, Charles Townes, comes in," says Genzel. He joined Townes's research group at the University of California, Berkeley, in the 1980s, where he helped to develop IR spectrometers for measuring the motions of gas clouds at different distances from Sgr A*. With those tools the researchers determined that clouds of ionized gas closer to the galactic center moved faster than those farther away, in a relationship that followed Kepler's laws of motion.

From those velocities, Townes and Genzel calculated that the gravity fields associated with a point mass of 4 million solar masses could drive the motion of the interstellar gases.[2] Such a massive object would be consistent with a black hole. "But nobody believed us," says Genzel. Other astronomers argued that winds or magnetic fields could also cause gases to move. Nonetheless, Genzel says, "we were fully convinced that it was a black hole." To convince the larger astronomy community, the Berkeley team realized, would require tracking the motions of individual stars.

Star of the show

By the early 1990s, near-IR speckle imaging was the best available tool for measuring the positions of distant stars. The technique produces clear composite images of celestial objects otherwise blurred by atmospheric turbulence. It does so via many short-exposure photographs, each of which captures a snapshot of the time-varying turbulence. Then the short exposures are spatially shifted to align the pattern of stars captured and are stacked on top of each other to provide a sharper image than any individual snapshot. (See the Quick Study by Steve Howell and Elliott Horch, PHYSICS TODAY, November 2018, page 78.)

Genzel and Ghez each focused their team's efforts on taking speckle images of the stellar cluster surrounding the galactic center. For three years Ghez and her colleagues at UCLA used the near-IR camera at the W. M. Keck Observatory in Hawaii to trace shifts in the positions of 90 stars across the sky.[3] Genzel and his colleagues at the Max Planck Institute for Extraterrestrial Physics in Germany used the 3.5-meter New Technology Telescope at the European Southern Observatory (ESO) in Chile to trace the positions of 39 stars, some shown in figure 1, over a four-year period.[4]

Both groups concluded by the late 1990s that our galaxy harbors a central dark mass of two to three million solar masses (see figure 1). "That work opened the door to showing that there is something massive in the region," says Tuan Do, a former student of Ghez and current deputy director of the galactic center group at UCLA.

According to Genzel, the initial stellar velocity measurements convinced astronomers that a black hole was present, but it did not convince physicists, whose skepticism originated in the reliability of far-away observations that didn't explicitly measure orbital accelerations. Both teams understood that to fully resolve the problem, they would need direct measurements not just of changes in the stars' positions but also of their radial velocities—that is, their motion along the observer's line of sight. But doing so was impossible with speckle imaging. Short exposure times limited monitoring to only the brightest stars. Velocity measurements required lengthy surveys and amounted to a marathon of beating down errors.

Fortunately, says Do, "this is also when technology started really improving." The new technology, known as adaptive optics, was first proposed in 1953. It allows Earth-bound telescopes to compensate for blurring due to atmospheric turbulence and see almost as clearly as if they were in space. The technique became available to Ghez's team at the Keck Observatory in 2000 and to Genzel's team at the ESO in 2003.

Adaptive optics relies on a bright, fixed reference object in the sky—either a bright known star near the observation target or an artificial star created by laser excitation of sodium atoms in the atmosphere. Sensors measure distortions in the wavefront of light a telescope receives from the reference object. Those distortion measurements enter a feedback loop that deforms a mirror in the telescope to correct for blurring. (See the article by Laird Thompson, PHYSICS TODAY, December 1994, page 24.) The technology also allows for the use of a spectrograph, an instrument that enables direct measurements of star compositions and radial velocities.

In 2002 Ghez and her colleagues completed the first measurements of an individual star, which they called S02, in the Milky Way's central stellar cluster.[5] The researchers monitored the Doppler shifts in S02's emission spectrum, inferred its velocity as it travelled around Sgr A*, and pinpointed the mass and the location of the object that it must be orbiting. "No assumptions other than Kepler's laws were required to determine the black hole mass and the distance from the star to the galactic center," says Do.

Genzel's team made similar observations to trace the orbit of the same star, which they called simply S2. The two teams' measurements aligned precisely with a highly elliptical orbit around a black hole of 4 million solar masses. The star approached Sgr A* at a perihelion distance of 17 light-hours.[6] "By the time

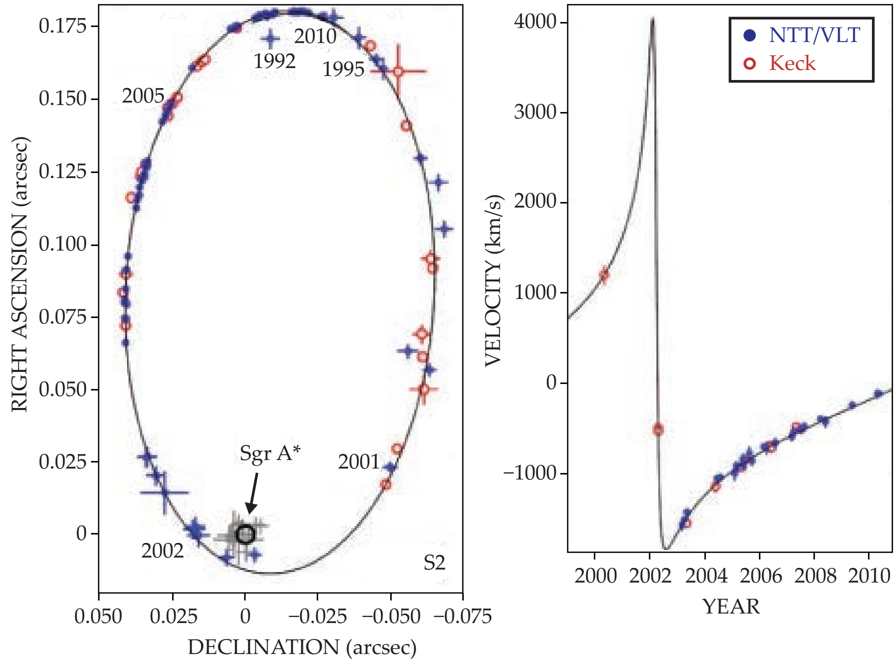


FIGURE 2. THE STAR S2 FOLLOWS AN ELLIPTICAL ORBIT AROUND SAGITTARIUS A*. The star's offset from Sgr A* as seen on the sky (left) and its radial velocity (right). The blue circles indicate data from the New Technology Telescope (NTT) and Very Large Telescope (VLT); the red circles are from the Keck Observatory. The gray crosses show the positions of IR flares that have been observed near Sgr A*. The data were collected through 2010. (Adapted from ref. 7.)

a single orbit was measured, physicists took it seriously," says Genzel.

Several more years of observations, illustrated in figure 2, continued to show excellent agreement between the Keck telescope and the ESO's Very Large Telescope. Analyses of combined data sets yielded refined values of every parameter that defined the orbit. "They showed independently that the mass at the center was so densely concentrated that no models besides a supermassive black hole should work," says Do. The observations ruled out other possibilities such as a spatially extended cluster of stellar-mass objects. Further support came from observations of near-IR flares from the location of Sgr A*, thought to indicate mass falling into a black hole.[7]

Precise precessions

Knowing that S2 should repeat its closest approach of the supermassive black hole in 2018, both teams eagerly awaited the opportunity to test a key prediction of Einstein's theory of general relativity: that an object's elliptical orbit should rotate gradually, or precess, at a rate different from that predicted by classical Newtonian mechanics. The prediction had been confirmed only in weak gravitational fields, such as the Sun's pull on Mercury. Testing it in a strong gravitational field became possible only with the observation of short-period stars in the galactic center. Both teams' measurements of the precession of S2's orbit showed that the theory does hold in strong gravitational fields.[8]

It would be surprising if the precession of stellar orbits strayed from general relativity, says Scott Tremaine of the Institute for Advanced Study in Princeton, New Jersey. However, as the measurement precision improves and observational time span increases, slight deviations from relativistic precession could reveal subtle constraints on theories of gravity that have not yet been quantified. Those deviations could also provide clues about other nonluminous matter—for example, stars too faint to be seen—that lies within a star's measured orbit. "The trick is to disentangle the effects of other stars from the effects of general relativity," says Tremaine.

That task requires examining the orbits of other stars close to the galactic center. The data set gathered by both Genzel and Ghez revealed other stars fainter than S2 that were at times hidden behind each other, but their orbits can be partially recovered by combing through older data (figure 3). Particularly exciting is the prospect of finding and measuring the orbits of stars that are even closer to the galactic center than S2. Those shorter-orbit stars should experience a stronger relativistic effect. "We don't know how close to the black hole you can get before stars die out. The hope is that the measurements will discover even closer stars," says Tremaine.

Emerging technologies will help. Researchers at the ESO recently developed the Very Large Telescope interferometer, which combines observations from four telescopes to generate images that have the spatial resolution equivalent to that of a 130 m telescope. Meanwhile, the Keck Observatory is planning a second-generation adaptive optics system.

Hopes and expectations

"Everything we've seen so far is consis-

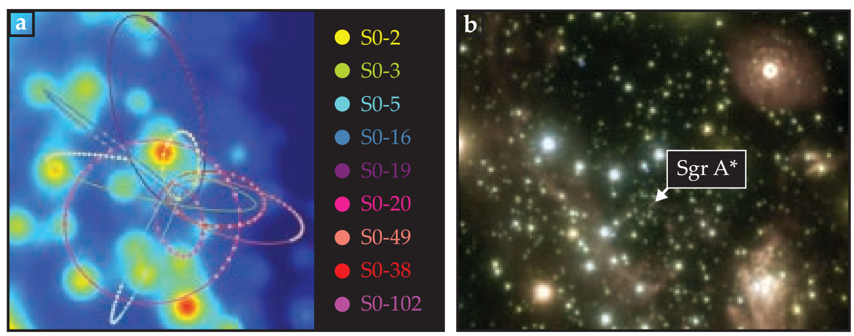


FIGURE 3. IMPROVED TECHNOLOGY HAS BROUGHT FINER-SCALE VIEWS OF THE GALACTIC CENTER. (a) Twenty-five years of observations have revealed the movements of many stars, including S02, within the central 1.0 × 1.0 arcseconds of the Milky Way. The annual average star positions are plotted as colored dots. Also shown are the best-fitting orbital solutions assuming a central supermassive black hole. The background is a 1.0 × 1.0 arcsec diffraction-limited image of the galactic center from 2015. **(b)** An image taken with the laser-guide-star adaptive-optics system at the Keck Observatory shows the central galaxy's central 10 × 10 arcsec region. (Courtesy of Keck Observatory/UCLA Galactic Center Group.)

tent with general relativity," says Genzel. Nevertheless, he notes, irrefutable proof that a candidate black hole is indeed a black hole requires determining values for the other specific parameters predicted by general relativity. In addition to mass, that includes the black hole's rotation rate. According to simulations, the orbit of another star at half to a third of S2's distance to the galactic center could provide the data necessary to calculate the spin of the proposed Milky Way black hole and its gravitational field. As the bright star S2 moves farther away from the galactic center, plotting the orbits of fainter, closer-in stars will become easier.

The view of the galactic center provided by Genzel and Ghez broadens the possibilities of answering complex questions. How does energy released during a black hole's formation influence the development of the rest of the galaxy? At what rate do stars fall into a central black hole? How does a black hole become supermassive?

Adaptive optics have already shown a surprising number of young stars and a lack of late-type stars near Sgr A*. "It's a hostile environment for making stars, so it stretches our understanding about star formation," says Tremaine. If stars get close enough to be tidally disrupted, they produce flares. Their interaction with the black hole may create hypervelocity stars that escape the galaxy extremely quickly. (See the article by Warren Brown, PHYSICS TODAY, June 2016, page 52.)

The radio-emitting object at the center of our galaxy made its debut as an intriguing source of "star noise" that interfered with terrestrial telecommunications. It now marks a region of the galaxy that can serve as an observational laboratory for exploring the impact of a supermassive black hole on stellar and interstellar environments. Shelley Wright (UC San Diego), who designs instrumentation for the Keck Observatory, says "the best physics that they're going to learn is still to come."

Rachel Berkowitz

References

1. B. Balick, R. L. Brown, *Astrophys. J.* **194**, 265 (1974).
2. M. K. Crawford et al., *Nature* **315**, 467 (1985).
3. A. M. Ghez et al., *Astrophys. J.* **509**, 678 (1998).
4. A. Eckart, R. Genzel, *Nature* **383**, 415 (1996).
5. A. M. Ghez et al., *Astrophys. J. Lett.* **586**, L127 (2003).
6. R. Schödel et al., *Nature* **419**, 694 (2002).
7. R. Genzel, F. Eisenhauer, S. Gillessen, *Rev. Mod. Phys.* **82**, 3121 (2010).
8. R. Abuter et al., *Astron. Astrophys.* **615**, L15 (2018); T. Do et al., *Science* **365**, 664 (2019).

Charpentier, Doudna win chemistry Nobel for development of CRISPR-Cas genome editing

Although less than a decade old, the technology is already shaking up research in cell biology and molecular biophysics.

Designer babies. Reviving the woolly mammoth. With prospective uses that sound more like science fiction than science fact, it's no wonder that CRISPR-Cas—the versatile biomolecular tool that gives researchers the power to make precise changes to cells' genetic codes—has captured the popular imagination.

But beyond the hype—and the thorny ethical questions—CRISPR-Cas is making waves throughout the biosciences. Inherited diseases whose genetic basis is well understood, such as cystic fibrosis and sickle cell anemia, could become routinely treatable. Agricultural crops could be more easily engineered with sturdier stems, better drought tolerance, and other desirable features. And across the basic-research community, CRISPR-Cas is revolutionizing what scientists can do with cells and DNA, even in contexts seemingly far removed from genome modification.

"I think of it as the equivalent of equipping your car with a GPS system," says Jie Xiao of Johns Hopkins University. "All you have to do is program in the coordinates and you can go anywhere in the genome that you want. You can bring lots of tools with you, and when you get there, you can do anything."

HALLBAUER & FIORETTI, BRAUNSCHWEIG, GERMANY

UNIVERSITY OF CALIFORNIA, BERKELEY

The discovery of CRISPR-Cas as a tool stems from a 2012 experiment by the groups of Emmanuelle Charpentier (then at Umeå University in Sweden, now at the Max Planck Unit for the Science of Pathogens in Berlin) and Jennifer Doudna (University of California, Berkeley).[1] Now the Royal Swedish Academy of Sciences has chosen Charpentier and Doudna for this year's Nobel Prize in Chemistry.

From bacteria to biotechnology

Although commonly called a gene edi-

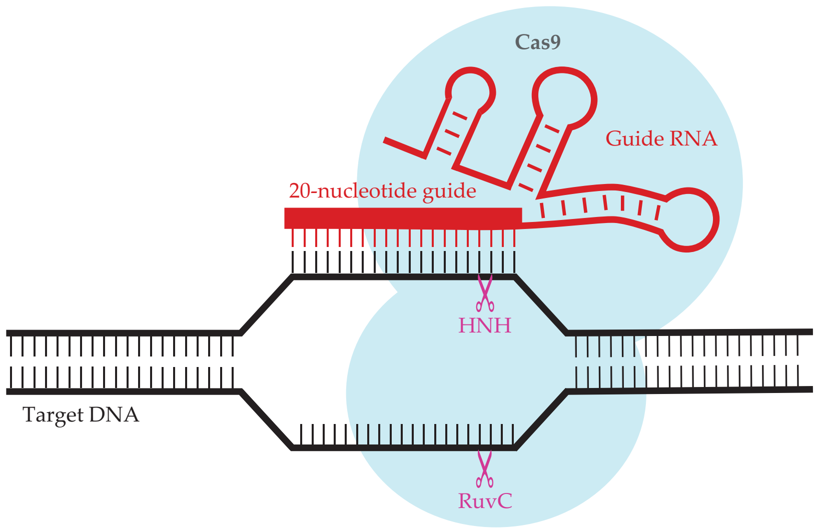


FIGURE 1. ADAPTED FROM A BACTERIAL IMMUNE SYSTEM, the CRISPR-Cas9 gene editor consists of the Cas9 protein (light blue) and a strand of guide RNA (red) that contains a 20-nucleotide guide sequence. The complex seeks out and binds to the complementary DNA sequence; two active sites (HNH and RuvC, purple) then cleave the DNA. (Adapted from ref. 5.)

tor, CRISPR-Cas is more precisely considered a gene cutter. In its canonical form, it seeks out a specific sequence of DNA and snips both strands of the double helix. To introduce the desired edits—to add, delete, or modify base pairs—researchers harness a cell's existing pathways for repairing its own DNA. The methods for doing so were already part of the pre-CRISPR-Cas genetic engineering toolbox.

The toolbox also contained two other gene cutters: zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs). Both those classes of proteins target DNA just like CRISPR-Cas does. The DNA sequence they target, however, requires an amino-acid sequence in the protein to recognize and interact with it. Getting that sequence just right—and getting the protein to fold into the right conformation—is a difficult and laborious process.

The beauty of CRISPR-Cas is that the target DNA sequence is simply the complement of an RNA guide sequence held by the Cas (short for CRISPR-associated) protein. Targeting a new gene is as easy as swapping out one piece of RNA for another.

The name CRISPR itself, which stands for "clustered regularly interspaced short palindromic repeats," has nothing to do with genome editing—and everything to do with where CRISPR-Cas systems were discovered in nature: as part of a microbial immune system. Viruses attack bacteria by injecting their genetic material into the single-celled organisms and commandeering their cellular machinery to make more viruses. To fend off those attacks, bacteria use CRISPR-Cas to cut the viral DNA and render it harmless.

To know which strands to cut, the bacteria store in their genomes short segments of DNA from every virus that they (or their ancestors) have ever encountered. Those viral snippets are separated from one another by a series of identical, symmetrical DNA sequences—the CRISPR.

The unusual repeating motifs in bacterial genomes were first observed in 1987, and their immune function was elucidated in the mid to late 2000s. A small but vibrant research field sprang up to identify and classify the various Cas proteins present in different bacterial species and to figure out how they worked in conjunction with other molecules to protect bacteria from viruses. The nascent discipline attracted both Charpentier, a microbiologist, and Doudna, a biochemist.

In their landmark 2012 discovery, Charpentier, Doudna, and their research groups built a simplified CRISPR-Cas system, outside of its native bacterium, that they could program with RNA of their choice to target and cut any DNA.[1] They chose the protein Cas9, from the bacterium *Streptococcus pyogenes*, for its all-in-one function: Unlike some bacterial species, which employ teams of several different Cas proteins to recognize and cut the target DNA, *S. pyogenes* uses just one.

A key insight was that the guide RNA, shown in red in figure 1, is considerably more complicated than the 20-nucleotide target sequence. *S. pyogenes*, in fact, uses two RNA molecules, both of which are essential to CRISPR-Cas9's function. But Charpentier and Doudna found that the two RNA pieces could be combined into one.

When Charpentier and Doudna demonstrated CRISPR-Cas9's programmable DNA-cutting activity *in vitro*, it still wasn't obvious that the system would work on the genomes of higher organisms. Plants and animals have a lot more DNA than either viruses or bacteria do, and they package it in dense, compact chromosomes inside a cell nucleus. As it turned out, however, none of those potential obstacles is enough to keep CRISPR-Cas9 from finding its target: Within just a few months of Charpentier and Doudna's publication, three other groups showed that the gene cutter can be made to work in human cells.[2]

On target

Viruses mutate over time. For CRISPR-equipped bacteria to mount an effective defense, their immune systems should be able to recognize viral DNA that isn't quite identical to the viruses they've encountered before. And indeed, wild-type CRISPR-Cas systems can cleave DNA that differs from their target sequence in two or three places.

But for editing human genes, it's important to get an exact match. The 3 billion base pairs in the human genome (counting just 1 of each of the 23 pairs of chromosomes) is small compared with the 1 trillion possible 20-nucleotide Cas9 target sequences. It's unlikely that a target sequence will occur in two or more unrelated places in the genome. But when CRISPR-Cas9 also targets near-miss sequences, the chances of it cutting the wrong gene are much greater.

Guided by theoretical and experimental studies of how CRISPR-Cas9 works, researchers have introduced mutations into the Cas9 structure to increase its specificity and limit the so-called off-target effects. (See the article by Giulia Palermo, Clarisse G. Ricci, and J. Andrew McCammon, PHYSICS TODAY, April 2019,

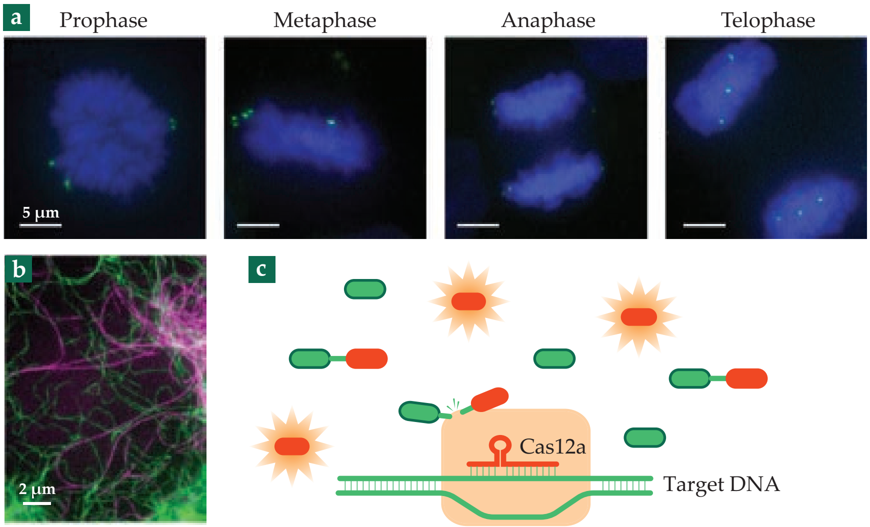


FIGURE 2. JUST A FEW OF CRISPR-CAS'S MANY APPLICATIONS. (a) Genes tagged with fluorescently labeled dCas9 (green) are imaged at different stages of mitosis. **(b)** The cytoskeletal protein vimentin (green) in a living cell is made fluorescent by modifying its gene with CRISPR-Cas9. (The protein tubulin, purple, is stained with a fluorescent dye.) **(c)** Target DNA from a virus is detected with Cas12a, which goes on to cut the DNA tethers holding reporter molecules together. Once cut, the reporter molecules become fluorescent. The detection scheme can be used to test for COVID-19. (Panel a adapted from ref. 6; panel b adapted from ref. 7; panel c by Freddie Pagani.)

page 30.) One approach has been to swap out some of the amino acids that bind to the backbone of the target DNA molecule—which looks the same from one DNA sequence to the next—to weaken that interaction and force the CRISPR-Cas9 complex to rely more heavily on binding between the guide RNA and its complementary sequence.[3]

But there's a delicate balance to be struck. CRISPR-Cas9 consumes no ATP or any other molecular source of energy. The energy it needs to unzip the target DNA must be supplied by the energy released when the complex binds to the target sequence, so if that interaction is made too weak, the system won't work. Ahmet Yildiz of the University of California, Berkeley, notes that although several versions of Cas9 with enhanced specificity have been developed, "it's still a challenge to avoid off-target effects without compromising the efficiency of on-target editing."

The gene cutter that doesn't cut

Tweaking Cas9's amino-acid sequence can not only refine the protein's activity but also change it. As figure 1 shows, Cas9 has two active sites for cutting the two strands of its target DNA. Deactivating one of them with a mutation creates a so-called nickase, an enzyme that cleaves just one of the DNA strands. And mutating them both creates dCas9—or dead Cas9—which binds to the target without cutting it at all. Perhaps counterintuitively, dCas9 is the vehicle for some of CRISPR-Cas's most promising applications.

The cuts and repairs induced by regular Cas9 are permanent, and that's too heavy a hand for some experiments on gene function. But when dCas9 binds, temporarily, to its target gene, it blocks the cell's transcription machinery from accessing that gene—in effect, switching it off. The ability to turn genes on and off at will lets researchers investigate how genes interact with one another to form circuits and how cells control which genes they express and which ones they don't. Almost every cell in the human body has all the same DNA, but the proteins in a muscle cell differ from those in a nerve cell. The capacity of cells to regulate their gene-expression profiles is essential to the development of complex multicellular organisms.

And dCas9 can do more to control gene activity besides physically blocking its target gene. Before sending dCas9 into the cell, researchers can affix to it one of the cell's own DNA-processing proteins, such as an activator or a repressor, so that when dCas9 binds to its target sequence, it boosts or restricts the expression of the gene next door. "Previous methods for doing this were not so straightforward," says Yildiz, "but with dCas9, we can block or control genes, change the gene expression profile, and switch from gene to gene without much trouble."

Furthermore, dCas9 can be decorated with fluorescent tags that enable researchers to image the location of the target sequence inside the nucleus—and ultimately the spatial arrangement of the whole genome. It's long been known that chromosomes' folded conformations are intricately hierarchical. But the leading method for probing them, called Hi-C, has its limitations.

In a Hi-C experiment, the cell is flooded with chemical links that bind DNA strands that are close in space; sequencing the DNA on either side of the links reveals which genes on which chromosomes are positioned together inside the cell. (See PHYSICS TODAY, December 2009, page 19.) That relative spatial arrangement is associated with genes' expression activity: "Genes that are near each other can share transcription enzymes," notes Xiao. But genes' absolute positions inside the nucleus, invisible to Hi-C, are important too. With fluorescent dCas9, researchers can visualize individual genes over time, as in the images in figure 2a of a cell undergoing mitosis.

Bringing cells into view

Imaging applications of CRISPR-Cas systems extend beyond looking at the genome itself. Fluorescence imaging can uncover hidden structures in all parts of the cell—and with superresolution techniques, it can do so in spectacular detail. (See PHYSICS TODAY, December 2014, page 18.) Most biomolecules, however, don't fluoresce on their own, so they need to be labeled with fluorescent tags.

For chemically fixed (that is, dead) cells, components can be labeled by staining the cell with fluorescent dye molecules bound to antibodies. The antibodies seek out and attach themselves to the protein molecules of interest, whose locations in the cell can then be imaged.

But dead cells aren't suitable for studying cellular dynamics. For imaging live ones, the usual technique is to force the cells to synthesize so-called fusion proteins: the protein of interest bound to either a fluorescent protein (see PHYSICS TODAY, December 2008, page 20) or a protein tag that attracts a fluorescent dye that's added later. The gene for the fusion protein is encoded on a plasmid—a small piece of DNA extraneous to the cell's own genetic code—and injected into the cell.

The plasmid-endowed cell has two versions of the protein gene: the one it carries naturally and the fusion gene on the plasmid. To maximize their chances of seeing the protein fluoresce, researchers typically induce the cell to overexpress the plasmid gene—that is, to make so many copies of the fusion protein that it overwhelms the cell's natural production of the untagged protein.

That overexpression, however, throws the cell's protein populations out of balance, and researchers can't be sure that the cell they're studying is representative of the cell's natural state. "If you have too much of the fusion protein and all its natural localizations in the cell are occupied, the excess proteins might go somewhere else," explains Stefan Jakobs of the Max Planck Institute for Biophysical Chemistry. "That may result in a wrong conclusion about the proteins' subcellular localization."

Genome editing solves that problem: If researchers can modify the cell's existing gene to include the fusion component, they can study the fusion protein expressed in natural amounts. They've been doing just that, with ZFNs and TALENs, before CRISPR-Cas technology came along, but CRISPR-Cas's programmability has sped the process, and it's now routine to obtain images such as the one in figure 2b in live cells. "It's really amazing how quickly it's been adopted," says Jakobs.

A world of possibilities

Much of the CRISPR-Cas research so far has used modifications of the Cas9 protein from *S. pyogenes*. But the microbial world is full of other Cas proteins, and even other versions of Cas9. Cas proteins are classified by their function and their evolutionary relationships, but even ones with the same name can have considerable structural differences. The Cas9 from *Staphylococcus aureus*, for example, has two cutting sites that resemble those of the *S. pyogenes* Cas9, but it's almost 25% smaller.

The diminutive size is important for some medical applications. "When it comes to treating blood diseases, we can extract bone marrow cells from the patient, do genome editing in a petri dish, then reintroduce them into the patient," says Dipali Sashital of Iowa State University. "But editing things like liver cells is a lot more challenging, and it requires a more complicated delivery mechanism." Benign viruses, stripped of their genetic material, can be loaded with the ingredients of a CRISPR-Cas system and deliver them to cells in the body, but only if the payload is not too large. The *S. aureus* Cas9 is small enough to fit, but the *S. pyogenes* Cas9 is not.

Despite the diversity of Cas9 proteins, they're an overall minority in nature. Many more microbial species employ so-called type I systems, which have a markedly different behavior: A complex of several proteins—Cas5, Cas7, and Cas8—recognizes the target sequence, then recruits another protein, Cas3, which not only cuts the target DNA but demolishes a large part of it.

"Evolutionarily, that probably makes more sense," says Sashital. In its function in a bacterial immune system, a CRISPR-Cas system needs to render viral DNA unusable. The sledgehammer of Cas3 is at least as effective in that role as is the scalpel of Cas9. Although type I systems aren't well suited for precise gene editing, their capacity to make large genome deletions has its uses—for example, in experiments that seek to identify which genes are essential to a microorganism's survival.[4]

One of CRISPR-Cas's most promising practical applications comes from two relatively new additions to the Cas toolkit, Cas12a and Cas13. Like Cas9, both are all-in-one proteins that detect the target genetic material—DNA for Cas12a, RNA for Cas13—and attack it, and both make precise cuts in their target sequences. But they don't stop there. Binding to the target sequence turns both proteins into indiscriminate genome-cutting machines, cleaving any other DNA or RNA that happens to be around, regardless of its sequence.

Because they can cut many strands of DNA or RNA with just one detection of the target sequence, Cas12a and Cas13 can function as biosensors. The concept is shown in figure 2c: Researchers mix Cas12a, a DNA sample that may or may not contain the target sequence, and a large number of reporter molecules—pairs of proteins linked by single strands of DNA. Cutting the links causes one of the proteins to fluoresce or change color, an easy visual cue that the target sequence is present.

In particular, that scheme can be used to detect viruses by their characteristic genetic material. Developed in 2018, the technique has been quickly pursued by companies that include Mammoth Biosciences in California (cofounded by Doudna) and Sherlock Biosciences in Massachusetts. With the outbreak of the COVID-19 pandemic, both companies have adapted their platforms to test for the SARS-CoV-2 virus. "Our test was the first FDA-approved CRISPR product," says James Collins, a cofounder of Sherlock Biosciences, "and we've contracted with Integrated DNA Technologies to produce a million tests a week."

The DNA strands in the reporter molecules are not part of the code of life. Rather, they serve as structural elements—a polymer tether that Cas12a happens to be able to cleave. Repurposed DNA has been explored for many other materials science applications, either by itself or integrated with other materials (see, for example, PHYSICS TODAY, April 2012, page 20, and "DNA origami assembles 3D arrays of nanomaterials," PHYSICS TODAY online, 7 February 2020).

Any number of DNA-based materials can now be made to respond to CRISPR-Cas systems in potentially useful ways. Before the pandemic, Collins and his group demonstrated a smart hydrogel made of DNA and other polymers. When Cas12a detected its target DNA sequence in the environment, it would cleave the DNA in the hydrogel to either release a cargo molecule or change the material's mechanical or electrical properties. "This line of research is just getting started," says Collins. "But a lot of researchers are wondering, 'What else can we use this for?'"

Johanna Miller

References

1. M. Jinek et al., *Science* **337**, 816 (2012).
2. L. Cong et al., *Science* **339**, 819 (2013); P. Mali et al., *Science* **339**, 823 (2013); S. W. Cho et al., *Nat. Biotechnol.* **31**, 230 (2013).
3. B. P. Kleinstiver et al., *Nature* **529**, 490 (2016); I. M. Slaymaker et al., *Science* **351**, 84 (2016).
4. See, for example, B. Csörgő et al., *Nat. Methods* (2020), doi:10.1038/s41592-020-00980-w.
5. F. Jiang, J. A. Doudna, *Annu. Rev. Biophys.* **46**, 505 (2017).
6. B. Chen et al., *Cell* **155**, 1479 (2013).
7. A. N. Butkevich et al., *ACS Chem. Biol.* **13**, 475 (2018).

PT

High radiation dose rates may improve cancer therapy

The ill-understood effect is gaining momentum and opening new avenues of research. CERN's Compact Linear Collider and other particle accelerators are contributing.

CAT PATIENTS SUFFERING from squamous cell carcinoma on their noses are treated with the standard of care—namely 10 roughly 1-minute-duration doses of 4.8 gray (Gy) each over five days (Muschti, top row)—or with FLASH radiotherapy in a single 30 Gy dose delivered in 20 ms (Lolo, bottom row). The ongoing trial is a collaboration between the Lausanne University Hospital and the Zurich Veterinary School, supported by the Swiss Cancer League. It is expected to enroll 29 cats total and be completed next year. From left to right, the cats are shown pretreatment, two weeks posttreatment, and six months posttreatment. (Photos courtesy of Carla Rohrer Bley and Marie-Catherine Vozenin.)

Dose rates hundreds to thousands of times higher than currently used in clinical treatment show promise for killing tumors while largely sparing healthy tissue. The mechanisms underlying the so-called FLASH effect are not understood, but results so far—in animals and in one human—have invigorated research into what some radiation oncologists and medical physicists say could revolutionize cancer treatment.

Scientists at medical and high-energy-physics research centers worldwide are tweaking existing accelerators to study and characterize the FLASH effect, named for the brevity of radiation exposure. The total dose is comparable to conventional therapy, but it is delivered in intense, subsecond single pulses or pulse sequences. Last month Varian launched a clinical trial that is treating bone metastases with proton FLASH therapy. Other companies are also getting involved.

In Switzerland, medical researchers at the Lausanne University Hospital (CHUV) are collaborating with physicists from the Compact Linear Collider (CLIC) project at CERN on a new accelerator to produce high-energy electrons for the first FLASH radiotherapy clinical trials on deep-seated tumors. They plan to begin the trials by 2024.

Studies are underway to better understand dose rates, why healthy tissue is spared, and more. "Think about the time response to radiation in tissue," says FLASH pioneer Marie-Catherine Vozenin, a radiation oncology professor at CHUV. "This involves novel biology questions, physics questions, chemistry questions. It's all crazy interesting." Radiation chemistry and radiation biology research were stagnant for a while, she adds. "That's not the case any more. The FLASH effect opens up work for the younger generation."

Even medical physicists who are more cautious about the potential of FLASH share a sense of excitement about the research it sparks. Daniel Low, a radiation oncology professor and vice chair of medical physics at UCLA, says the bar is high for widespread change in therapy protocols. "How hard will it be to get to parity with conventional radiotherapy? Well beyond parity?" If the FLASH effect of sparing normal tissue while damaging tumors works for cancers in humans, he says, "you can be sloppier in delivery because you can take advantage of healthy tissue. But we don't yet know the tradeoffs." And because of necessary financial investment, any new approach "needs to knock it out of the park." Although he doubts that FLASH will "step up," Low notes that radiotherapy research in the past 10 years or so has been about "dotting i's and crossing t's. This is much more exciting."

Unanswered questions

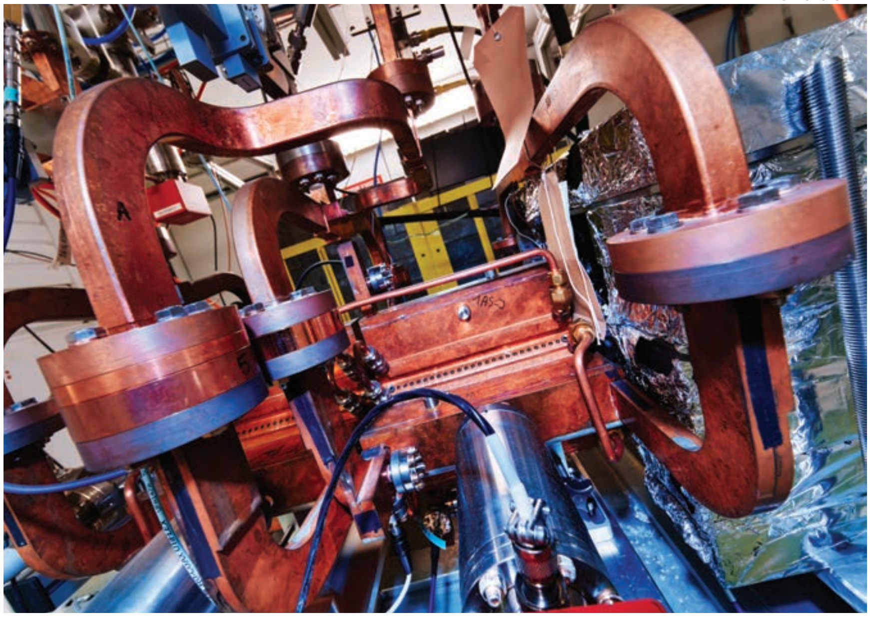
The radioprotective effect of ultrahigh dose rates was observed half a century ago on cells and bacteria in Petri dishes. Work was translated to animal studies but not to clinical ones, and research into high dose rates fizzled out. About 13 years ago, Vozenin and her colleague Vincent Favaudon used an experimental linear electron accelerator in Paris that could operate at both conventional dose rates of less than 0.03 Gy/s and at ultrahigh dose rates, in excess of 40 Gy/s. (The SI unit of ionizing radiation dose is the gray, defined as the absorption of one joule of radiation energy per kilogram of matter.)

"Out of scientific curiosity," Vozenin says, they exposed mice with lung tumors to radiation at high dose rates. "We had access to the beam, and we thought it had never been investigated." They were surprised, she says, that delivering the same integrated dose as in conventional radiotherapy retained the cure while all but eliminating the toxicity to surrounding tissue. In a 2014 paper reporting the findings, the team concluded that "FLASH radiotherapy might allow complete eradication of lung tumors and reduce the occurrence and severity of early and late complications affecting normal tissue." Subsequent studies have demonstrated the FLASH effect in pigs, rats, cats, zebrafish, and other animals.

A popular hypothesis for why FLASH radiotherapy spares healthy tissue has to do with molecular oxygen concentration. Cells are killed by mechanisms that include DNA breakage, lipid damage, and changes in cytokine cascades and inflammation signals, all of which are affected by the availability of reactive oxygen species. Solid tumors are typically hypoxic, and their response to radiation appears to be independent of dose rate. And healthy tissue appears to gain protection from radiation damage by the reduced production of free radicals in FLASH compared with conventional radiation. "FLASH exploits the difference in cell metabolism between normal and tumor tissue," Vozenin says.

But the oxygen hypothesis doesn't fully explain the FLASH effect. For example, Emil Schüler, a medical physicist at MD Anderson Cancer Center in Houston, Texas, notes that the mechanism of tumor destruction with proton radiation is less dependent than photon and electron radiation on oxygen concentration. Even so, a FLASH effect with protons still occurs. "We need to study the parameters," he says, "the mean dose rate, total time of delivery, dose per pulse, variations among organs, and more." Keeping tabs on dose delivery is also a concern. Most radiotherapy systems rely on ionization chambers for such monitoring, but they don't work properly at the subsecond time scales of FLASH. Other methods for measuring and standardizing the dose delivery are needed, he says.

© 2020 CERN



A HIGH-ENERGY FLASH TEST FACILITY to study the treatment of deep tumors is planned for the Lausanne University Hospital campus. The facility, which is expected to start up by 2024, is based on CERN's Compact Linear Collider. In this prototype, the electron beam travels from left to right and gains 25 MeV over the 25 cm length of the accelerating structure (center, copper with holes). The protruding pipes with rectangular cross section are waveguides that power the accelerating structure.

There are many unanswered questions, notes Peter Maxim, vice chair of the medical physics division at Indiana University School of Medicine. "I am convinced that the FLASH effect is real based on the published preclinical studies," he says. "The million-dollar question is whether the FLASH effect is translatable to human therapy." Untangling the mechanisms of FLASH isn't necessary for its use, he adds, but it would help in designing clinical trials. And, he notes, technological advances would be needed to implement FLASH for cancer treatment.

FLASH facts

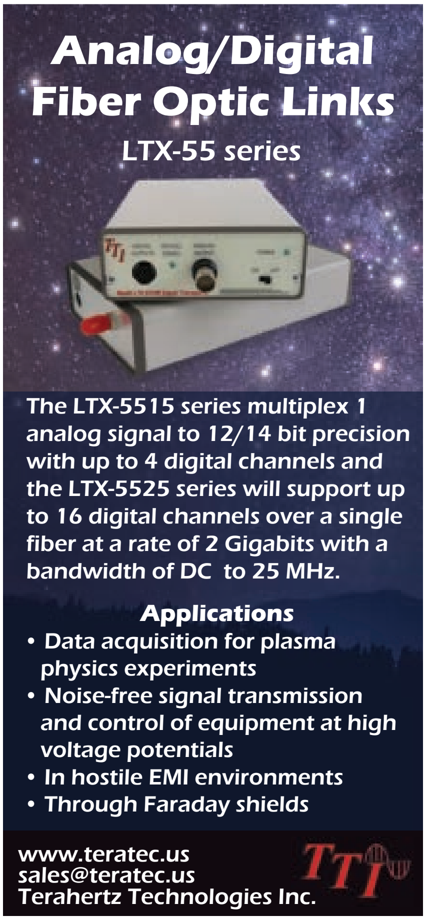
In FLASH radiotherapy studies, mean dose rates of at least 40 Gy/s are delivered in pulses in the microsecond to millisecond range; the details depend on the type of radiation and the accelerator. So far, most research has been done with electron beams, but protons and photons are also being studied, and scientists have observed the FLASH effect for all forms of ionizing radiation.

An advantage of electrons is that they can be readily produced at high fluxes. Photons are generated by bombarding electrons on a tungsten or other target, so to get high-energy photons, the electron energy has to be correspondingly higher. Protons are more expensive to accelerate, but they have advantages, one being that they can more accurately target tumor volumes. (See the article by Michael Goitein, Tony Lomax, and Eros Pedroni, in PHYSICS TODAY, September 2002, page 45, and the news story, June 2015, page 24.)

At TRIUMF, Canada's national particle accelerator center in Vancouver, researchers are gearing up to investigate FLASH with protons and photons. The most common radiation therapy at hospitals and clinics uses photons, says Cornelia Hoehr, who coleads the center's FLASH work: "There is more data, more protocols, it is mainstream, and it would be approved for clinical use faster."

Among the attractions of FLASH is that in delivering the radiation on subsecond time scales, organs and tumors are effectively motionless. In a single breath, for example, lungs can move as much as 2 cm, and movement in the bowels can push the prostate by 1 cm in a minute. On the subsecond time scale, such motions are negligible, and methods for simultaneous imaging and accurately targeting tumors are being developed. Because of the high dose rate that FLASH would deliver, the necessity for positional accuracy is enhanced, both in terms of killing cancer cells and minimizing harm to healthy ones.

Conventional radiation is typically delivered in 20–30 doses, each a few minutes in duration, over several weeks to



minimize damage to healthy tissue. In FLASH, the same total dose would be delivered in a single or just a few sittings, in milliseconds instead of minutes total. Even if the tissue sparing doesn't hold for humans, reducing clinic visits would be a huge improvement in care, says Hoehr. Fewer visits would also reduce the cost of treatment.

And given that radiation therapy is in part guided by the tolerance of healthy tissue, FLASH may make it possible to give higher doses and therefore treat more types of tumors or treat them more effectively. About half of cancer patients are treated with radiation, usually in conjunction with other treatments.

Existing accelerators are limited by electron energies of 6–20 MeV to shallow tumors, down to about 4 cm. "If we can show that [FLASH] spares normal tissue, we cannot defend only using conventional dose rates anymore," says Julianne Pollard-Larkin, a medical physicist at MD Anderson. She notes that according to some models, FLASH treatment can reduce damage to healthy tissue by 70% compared to conventional radiation treatment. "FLASH looks like a win for science, for cancer patients, and for humanity."

High energies for deep tumors

For wider applicability and the potential to treat deep internal tumors, higher energies are necessary. The CERN–CHUV plan is to produce electrons in the 100–200 MeV range. "As soon as we learned of FLASH, my colleagues and I realized that CLIC technology could be relevant," says Walter Wuensch, a senior researcher at CERN. A future CLIC accelerator, if built, would use room-temperature, high-gradient technology to accelerate electrons and positrons. "The same features that make the high-energy collider not too huge or expensive also make for a good fit in a hospital campus," says Wuensch.

The building blocks adapted from CLIC for the FLASH therapy facility include high-gradient accelerating structures, suppression of beam instabilities, and a high-current photo-injector gun. Tests and simulations show that the design "will allow us to stably accelerate a lot of charge to get a high enough dose in a short time," says Wuensch. The beam will be accelerated to treatment energies in a bit more than a meter, he adds.

The Lausanne FLASH facility will use high-power microwave klystrons, which CLIC uses in tests but not in its final design. And specifications such as beam intensity, steering, and beam size are determined by the medical trial needs. The systems are similar, says Wuensch, "but to have hardware work with the reliability to treat human beings will cause some sleepless nights."

Jean Bourhis, head of CHUV radiation oncology, leads the medical side of the collaboration. He says the facility, including a building to house it, will cost around 30 million Swiss francs ($33 million). CHUV is raising funds for the facility, he says. He expects the new building and the machine to be ready by 2023 and clinical studies on deep-seated tumors to start soon thereafter. The facility is designed such that trial parameters can be varied, and it will be larger than what a future hospital machine might need. For wide adoption, a facility would have to be less variable and easier to use.

In the meantime, CHUV is conducting clinical trials on superficial tumors. "We are monitoring toxicity and tumor control and comparing FLASH to conventional radiation therapy for skin cancer," says Bourhis. In 2018 the CHUV team treated a single patient with FLASH with excellent results, he says. The patient had undergone conventional radiation therapy more than 100 times and suffered continued discomfort. A single FLASH dose cleared a lesion completely, as of five months posttherapy.

For CLIC physicists, the collaboration with CHUV is inspiring both for advancing medicine and having societal impact and for the attention it brings their accelerator project, says Wuensch. CLIC has been something of a stepchild at CERN for more than two decades, a neat project that gets ongoing research funding but no green light. In the recent European Strategy for Particle Physics (see PHYSICS TODAY, September 2020, page 26), the Future Circular Collider is a high priority, but if it proves too expensive or otherwise infeasible, plan B is CLIC.

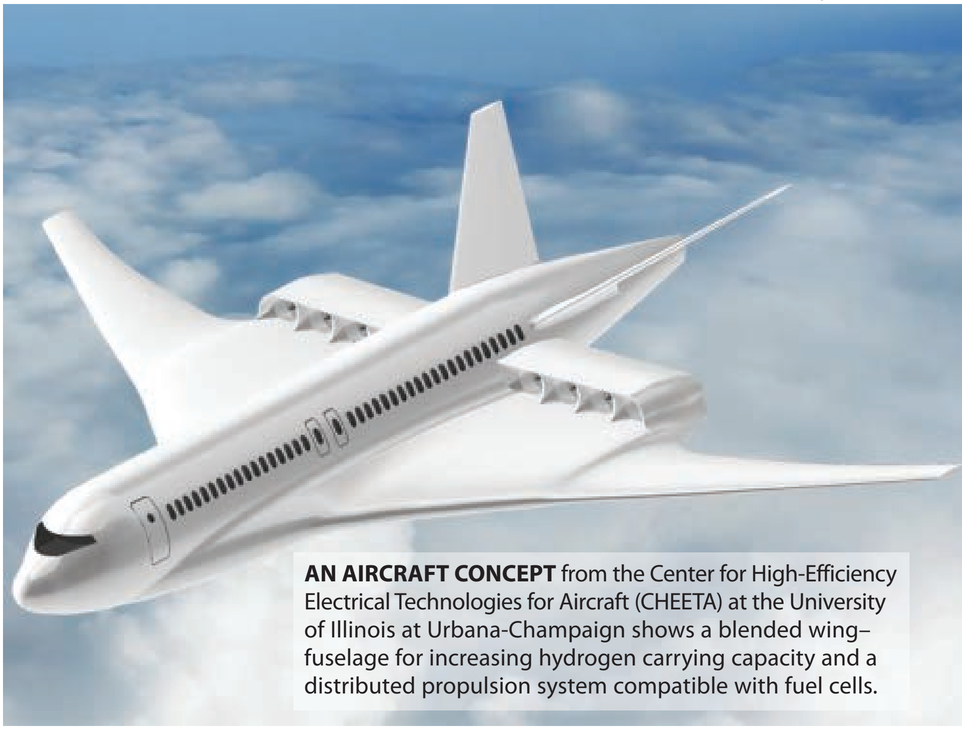
The collaboration with CHUV, Wuensch says, could feed back into increased support for CLIC. What's more, he says, the collaboration "keeps and expands expertise in CLIC technology, and if everything goes well, clinical FLASH facilities could go into commercial production."

Toni Feder

Hydrogen-powered aircraft may be getting a lift

Cutting the weight of fuel tanks and continuing advances in fuel-cell technologies are key to making hydrogen competitive in aviation.

CHEETA



AN AIRCRAFT CONCEPT from the Center for High-Efficiency Electrical Technologies for Aircraft (CHEETA) at the University of Illinois at Urbana-Champaign shows a blended wing–fuselage for increasing hydrogen carrying capacity and a distributed propulsion system compatible with fuel cells.

As nations work toward achieving net-zero carbon economies, commercial aviation will be one of the most difficult sectors to decarbonize. Fossil fuel's energy-density advantage is too tough to beat, the argument goes, and rather than try to confront that, it might make more sense to continue the use of petroleum in airliners and to aim at offsetting their emissions using some negative carbon emissions technology.

Nonetheless, plenty of research is underway on lower-carbon aircraft propulsion. Some early-adopter airlines routinely blend biofuels into their aviation fuel, for example, and in 2018 Boeing flew a commercial airliner on 100% biofuel for the first time. But biofuels have their own drawbacks: The growing, gathering, and conversion of crops to liquid fuels is carbon-intensive. And biofuels' potential is limited by agricultural and other competing land uses.

Another option is hydrogen-powered flight. To be carbon-neutral, the hydrogen must be produced either with renewable energy or with natural gas equipped with carbon capture and storage. Both of the world's major airliner manufacturers are looking at the lightest element as one option for reducing their customers' carbon footprint. Amanda Simpson, vice president for research and technology at Airbus Americas, says Airbus will decide by 2025 whether the market can support hydrogen-fueled airliners. Assuming that it can, the company projects its first hydrogen airliners will enter service in 2035.

In 2008 Boeing built and operated the first aircraft ever to fly solely on hydrogen power. The fuel cells on the single-person plane were supplemented with power from lithium ion batteries during takeoff and ascent. Four years later, the company unveiled the Phantom Eye, a liquid-hydrogen-powered unmanned aerial vehicle (UAV). It was designed to fly reconnaissance missions of up to four days at an altitude of 20 000 meters. Boeing was unable to sell the UAV to the military, however, and it is now a museum piece.

Chris Raymond, Boeing's chief sustainability officer, says the company is looking at hydrogen for commercial aviation, but notes that the more immediate focus is on sustainable aviation fuels. "The reality is we have to do something now. But our view is that there will be a mix of solutions," with hydrogen power more likely to fill the short-haul, smaller end of the sector.

Although Boeing has shown that hydrogen will work as aviation fuel, more effort is required to determine whether an aircraft's structure and fuel tanks can be built to operate as safely as today's airliners, says Mike Sinnett, vice president of product development for Boeing's commercial airplanes division. He estimates it will be two decades or more before hydrogen could be introduced in Boeing aircraft, and he says engines that will power aircraft a decade from now are being designed today.

The small company ZeroAvia has set its sights on manufacturing a 10- to 20-passenger aircraft powered by hydrogen fuel cells by 2023. The California startup, according to its founder and CEO Val Miftakhov, has received $5 million in grants from three UK government programs and has attracted interest from 12 regional carriers in the UK, US, and European Union.

Weight and energy density

The biggest impediment to hydrogen-powered flight is the extra weight required for fuel storage, be it in gaseous or liquid form. For liquid hydrogen, the challenge will be making lightweight vacuum-insulated tanks that maintain the fuel below its 20 K boiling point. Gas carries a greater weight penalty, since the tanks must be built to withstand high pressures of 250–350 bar.

Measured by megajoules per kilogram, liquid hydrogen offers 2.8 times the energy density of aviation fuel. But in terms of combined fuel and tank weight, aviation fuel has the advantage over hydrogen by a factor of 1.6, says Rajesh Ahluwalia, a hydrogen and fuel-cell researcher at Argonne National Laboratory. Whereas aviation fuel constitutes about 78% of the combined weight of tank and fuel, liquid hydrogen accounts for just 18% of the total in current storage designs. To compete with fossil fuels, the fuel weight fraction has to reach at least 28%, he says.

Compared to hydrocarbons, liquid hydrogen has a much lower energy per unit volume. But Phillip Ansell, director of the NASA-funded Center for High-Efficiency Electrical Technologies for Aircraft at the University of Illinois at Urbana-Champaign, says that the added external surface area required to accommodate larger hydrogen tanks, and the resulting increase in aerodynamic drag, can be minimized by carefully tailoring other parts of the aircraft. One potential solution is a blended wing–fuselage design (see illustration, page 27).

ZeroAvia will use hydrogen compressed to 350 bar on the fuel-cell-powered aircraft that it plans to fly on short hops of 500 nautical miles (roughly 900 km) or less. Gaseous hydrogen could also find a niche in unmanned aerial vehicles, Ahluwalia says. But for large airliners carrying 150 or more people on longer flights, "gaseous hydrogen won't cut it," says Ansell. A pressurized storage system requires a much more robust tank and would occupy about twice as much space as tanks containing liquid hydrogen, he says.

By the end of the decade, ZeroAvia expects to debut 50- to 100-passenger fuel-cell aircraft powered by liquid hydrogen. The tanks would be made from combinations of existing composites and resins. No new materials development is needed, Miftakhov says. ZeroAvia's current fuel-to-tank weight ratio is 11–12% for gaseous hydrogen. The company is currently testing liquid hydrogen tanks that have ratios greater than 50%, he says.

Fuel cells or combustion?

Subsonic jetliners such as the Boeing 737 are powered by turbofans, which use mechanical energy derived from a gas turbine to accelerate air rearward by means of a ducted fan. Today's gas turbines could burn hydrogen with relatively few modifications. "You can almost just drop hydrogen into today's engines," says Ansell. Simpson compares the conversion process to adapting a propane grill to operate with natural gas. But though it would produce no CO_2 or soot, hydrogen will produce pollutant nitrogen oxides (due to nitrogen's presence in air) and water vapor. A greenhouse gas, water vapor at high altitudes remains in the atmosphere longer than at lower elevations, where it precipitates as rain.

Hydrogen-fueled proton exchange membrane (PEM) fuel cells are emissions-free if the hydrogen comes from carbon-free sources, and their exhaust water can be condensed before release. PEMs, though, provide only half the 3.7 kW/kg power per unit weight of modern gas turbines burning conventional fuel, says Ansell, not including the weight of the fuel or the tank. Still, that's an order of magnitude improvement from the 0.3 kW/kg of 15 years ago, and continued advances are likely.

Today's PEM fuel cells should compete with piston aircraft engines in powering four- to six-passenger aircraft, says Ahluwalia. But their energy-to-weight ratio is far below that of turboprops and turbines.

The fan of a turbofan provides about 80% of the engine's total thrust, with the remainder delivered by combustion. It's hoped that fuel-cell systems can be developed to deliver the entire thrust of a turbofan by electric power. Alternatively, fuel cells could be supplemented with battery power during takeoff and ascent.

"We'll probably never compete with the specific power of a gas turbine," says Dave Tew, a program manager for electric flight systems at the Department of Energy's Advanced Research Projects Agency–Energy (ARPA–E). But the hope is that the efficiency of fuel cells and electric drive systems can be improved enough to match or exceed the energy-to-weight ratios of the turbine and fuel combined.

Airbus has produced three concepts for hydrogen-fueled airliners with capacities of up to 200 passengers and ranges of 2000 nautical miles (3700 km) or more. Each is proposed to be powered by a hybrid system of combustion turbines and fuel-cell-driven motors. In a turboelectric configuration, a hydrogen-fueled gas turbine drives an electric generator, and the fan is driven by an electric motor.

"We think that at the 1000- to 2000-mile ranges, turbines will be required, but we will see as we go through the pencil sharp-

ZEROAVIA

ZEROAVIA's six-seat fuel-cell-powered aircraft made its maiden flight in September 2020 in Cranfield, UK. The company claims it is the world's largest hydrogen-powered aircraft.

AIRBUS

THREE CONCEPTUAL HYDROGEN-FUELED airliners from Airbus all combine hydrogen-burning gas turbines with hydrogen fuel cells that generate electrical power. The turboprop (bottom) would carry up to 100 passengers and have a range of 1000 nautical miles. The others would carry up to 200 passengers, with a range of 2000 nautical miles.

ening," Simpson says. "We've already seen some countries in Europe saying that aircraft within their borders will have to be zero [carbon] emission. Whether [those requirements] will be enough to launch a larger aircraft like we are talking about remains to be seen in interest from our customers," she says.

Airbus envisions synthetic fuels produced from renewable sources, not hydrogen, powering its future long-range 300- to 400-seat airliners. Though the company already is familiar with hydrogen in aerospace applications, adapting the fuel to aircraft that will be emptied and filled multiple times a day will be a new challenge, Simpson says.

R&D needed

Whereas PEM fuel cells rely on a liquid electrolyte to shuttle protons between electrodes, solid oxide fuel cells (SOFCs) use an oxide material, usually yttria-stabilized zirconia, as the electrolyte to produce electricity by oxidizing a fuel. The cells can be fueled by liquid hydrocarbons or ammonia. SOFCs offer higher efficiencies than PEMs. But they weigh significantly more, and heavier hydrocarbon fuels such as gasoline or kerosene must first be reacted by steam or air to form simpler molecules such as hydrogen and methane. For very long flights, the higher efficiency of SOFCs may make up for their greater weight, says Ansell.

In August ARPA–E awarded \$13.1 million in grants to six companies to develop different SOFC technologies for aviation. Tew says weight reduction is a major thrust of the R&D. Other DOE programs support PEM research, so ARPA–E's program specifically excluded them.

Thermal management at altitude will be an issue for any type of fuel cell, notes Ansell, since the lower air density diminishes capacity to shed waste heat. "The question is how efficiently you can reject the heat, and SOFCs are actually quite good at this." The high operating temperatures of SOFCs—around 800 °C, compared with PEMs' 100 °C—actually improve heat removal due to the wider differential between fuel-cell and ambient temperatures. Moreover, the waste heat from an SOFC might be directed to drive a gas turbine, in a manner analogous to an automobile turbocharger, he says.

To obtain oxygen, fuel cells require a sufficient airflow that may not be achievable at high cruising altitudes. That problem could be overcome by compressing air, albeit at the expense of power available for propulsion.

Hydrogen-fueled aircraft will need to meet the same levels of safety and integrity as those fueled with kerosene. "How do you inert a fuel tank with hydrogen?" says Boeing's Sinnett, using industry language for reducing the risk of fuel-tank flammability. "How do you fuel the airplane? How do you avoid electrostatics in the fueling process? How do you ensure the structural integrity of the fuel tank?" If solutions can be found to those questions, the next step is determining whether that aircraft can be efficient. "Almost anything is achievable from an engineering and technology standpoint, but you have to ask if the result will be practical," he says. Finally, the plane must be built to hydrogen-specific safety and regulatory standards that don't yet exist.

The infrastructure

Adoption of hydrogen fuel will be impossible without a fueling infrastructure. But the challenge should be less daunting for airlines than it is for car owners. If demand warranted, hydrogen could be produced on site, thereby eliminating distribution costs. ZeroAvia plans to do just that, using nearby renewable energy sources to power electrolyzers. In the company's scheme, a single airport hub could service multiple round-trip routes of its aircraft, so long as flights are 250 nautical miles (about 460 km) or less each way, Miftakhov says. That's suitable for flights between London and Paris or between New York and Boston.

Miftakhov predicts that the cost of fuel cells and electrolyzers should follow a similar trajectory as photovoltaics, whose costs have fallen more than 80% in the past 10 years. "In three years we will get to half of today's cost for hydrogen at the point of consumption, partly through improvements in scale." Even small airports could produce and dispense multiple tons of hydrogen per day, he says; a typical vehicle fueling station, for comparison, sells 200 kg a day. The high aviation volume should drive down the price from its uncompetitive level of \$12–\$15/kg at filling stations. The efficiency of electrolyzers also continues to improve markedly (see PHYSICS TODAY, August 2020, page 20).

Several interested parties have already approached Airbus about providing fuel to airports should the company decide to move forward on hydrogen, says Simpson. "The International Energy Agency has said we can expect to see green hydrogen capacity of 40 gigawatts in Europe by 2030, and we expect the rest of the world will have to jump on."

David Kramer PT

What a year this week has been!

Those words of weariness appear on a sign planted this Spring in the front yard of a house in my Capitol Hill neighborhood. It remains on display.

My colleagues and I at *Physics Today* have been less despondent during the current pandemic. By the time you read this, we'll have produced and published eight issues entirely from home. The magazine's website has likewise maintained its stream of prime content.

Keeping busy has kept us engaged. But there's another source of optimism that sustains us. Physicists have returned to their labs. We're starting to see papers appear in *Nature, Physical Review Letters,* and other journals that were submitted after shutdowns and lockdowns began.

Government support for research funding in the US and other countries remains robust. Congress is on track to provide NSF with at least a 2% budget increase for this fiscal year.

Research output receded during the pandemic, but it did not disappear, and now it is recovering. For your business before, during, and—we hope—after COVID-19, *Physics Today* thanks you!

Charles Day,
Editor in Chief

Accurion GmbH
Aerotech Inc.
AGICO, Inc.
American Astronomical Society
American Elements
Anderson Materials Evaluation, Inc.
Angstrom Sciences, Inc.
APE Angewandte Physik & Elektronik GmbH
AR RF/Microwave Instrumentation
Astro Met, Inc.
Avtech Electrosystems Ltd.
Bartington Instruments Ltd
Bristol Instruments, Inc.
CAEN Technologies, Inc.
California Institute of Technology
Cambridge Univ Press (UK)
Chicago Quantum Exchange
Circuit Insights LLC
ColdEdge Technologies Inc.
COMSOL INC.
Cotronics Corp.
Cremat Inc
Cryogenic Control Systems, Inc.
Cryogenic Limited
Cryomagnetics, Inc.
Cryomech Inc.
CS CLEAN SOLUTIONS
Duniway Stockroom Corp.
École Polytechnique Fédérale de Lausanne (EPFL)
Edwards Vacuum
Electrical Energy Limited, LLC
Florida State University
Ho-Am Foundation
The Hong Kong University of Science and Technology
HÜBNER Photonics
ICEoxford Ltd
IEEE
Innovative Sensor Technology, USA Division
Insaco Inc.
InstruTech
Integrated Magnetics
Intermag 2020
International Centre for Diffraction Data
Intlvac Thin Film Corporation
IOP Publishing
J.A. Woollam Co., Inc.
Janis Research - A Lake Shore Company
Karlsruhe Institute of Technology (KIT)
The Kavli Foundation
KEK, The High Energy Accelerator Research Organization
KEYCOM Corp.
Keysight Technologies
Kimball Physics
KNF Neuberger Inc.
Korea Institute for Advanced Study
Krell Institute
Lake Shore Cryotronics, Inc.

to our 2020 Industry Partners

Liquid Instruments
Los Alamos National Laboratory
Lumedica
Mad City Labs, Inc
Master Bond Inc.
Materials Research Society
The MathWorks, Inc.
Max Planck Institute for the Physics of Complex Systems
McPherson
MDC Precision Products LLC
MeiVac, Inc
Microscopy & Microanalysis (M&M) Conference
MIT
MLD Technologies, LLC
MMM 2020
Montana Instruments Corporation
Naieel Technology
NanoMagnetics Instruments Ltd.
The National Academies of Sciences, Engineering, and Medicine
National Academies Press
National Institute of Standards and Technology (NIST)
National Reconnaissance Office
NEXTRON
Nor-Cal Products, Inc.
North Star High Voltage
Oakland University
OriginLab Corporation
OSA, The Optical Society
Oxford Instruments
Pacific Northwest National Laboratory
Park Systems
Paul Scherrer Institut (PSI)
PDE Solutions Inc
Pearl Companies
Pearson Electronics
Pfeiffer Vacuum Inc.
Photon etc.
PI USA
PicoQuant GmbH
Pittcon
Polytechnique Montréal
Princeton Center for Complex Materials (PCCM)
Princeton Univiversity Press
Purdue University
Q-CTRL
Quantum Design Inc.
RD MATHIS COMPANY
Rohde & Schwarz GmbH & Co. KG
Saint-Gobain
SEMI
SHI Cryogenics Group
SPIE
Stanford Research Systems
STAR Cryoelectronics
SurfaceNet GmbH
Sutter Instrument Company
SVC
Synergy Software
Tabor Electronics Ltd.
Tech-X
tectra GmbH
TEKTRONIX, INC.
Terahertz Technologies Inc.
Thyracont Vacuum Instruments GmbH
TOPTICA Photonics Inc.
Trinity College
Universitat Heidelberg
The University of Chicago
University of Geneva
University of Notre Dame
The University of Mississippi
The University of Tennessee
The University of Texas at Austin
University of Wisconsin-Madison
University of Waterloo
Velmex, Inc.
Vistec Electron Beam GmbH
Wavelength Electronics, Inc.
Woods Hole Oceanographic Institution
XENON Corporation
Zurich Instruments AG

Exploring cultural heritage through ACOUSTIC DIGITAL RECONSTRUCTIONS

Representation of a musician inside a virtual-reality room. As a real musician plays, he is transported to the Château de Versailles, whose great hall is shown in the background and whose acoustics are "reconstructed" in real time. One microphone captures the instrument's sound while the musician's position and orientation are tracked in space. The audio is rendered either over headphones or through an array of loudspeakers. (From the EVAA project; see http://evaa.pasthasears.eu; image by David Poirier-Quinot.)

Brian Katz is a CNRS research director at the Sorbonne University, Institute d'Alembert in Paris, France. **Damian Murphy** is a professor of sound and music computing in the department of electronic engineering at the University of York in the UK. **Angelo Farina** is a professor in the department of engineering and architecture at the University of Parma in Italy.

Brian F. G. Katz, Damian Murphy, and Angelo Farina

Simulating the acoustics of destroyed or altered amphitheaters, cathedrals, and other architectural sites re-creates their sonic grandeur.

The fire at Notre Dame Cathedral in Paris in 2019 and the one at Gran Teatro La Fenice opera hall in Venice in 1996 are reminders of the fragile nature of humanity's cultural heritage. Fortunately, acoustic measurements, numerical simulations, and digital reconstructions can recover—and to some extent preserve—the sound of humanity's great architectural sites. What's more, those techniques provide a way for archaeologists, historians, musicologists, and the general public to experience the lost acoustics of damaged or destroyed places.

A room's sound

The objective of architectural acoustics is to achieve the best sound quality possible in a space, whether it's a theater, church, concert hall, or recording studio. The propagation of sound is subject to several factors. We speak of direct sound to represent the propagation path of a sound that reaches listeners without any obstacles in its way. Indoors, the presence of walls changes the direction of the acoustic energy. The new sound paths correspond to different distances and interactions with the architecture.

When a source emits a sound, the result is direct sound and reflections that are picked up by a receiver. The collection of those reflections over time constitutes the room's acoustic response. When the source stops producing, listeners perceive the sound's gradual decay as reverberation—the time it takes for the sound to fade away.

Before the formal theory of room acoustics was developed, the ancient Greeks put their experiential knowledge of sound into practice. Amphitheaters, such as the archaeological site of Tindari, Sicily, shown in figure 1, are representations of that work. If a roof and surrounding walls were added to an open-air Greek theater, the effect on listeners would be striking: The acoustic energy would be directed downward but dispersed in time. Reflected sounds would take different paths in the room and reach our ears at different times.

The acoustic quality of a room therefore depends, to first approximation, on the reverberation time, which can vary depending on the room's construction and decoration materials, the position of the sound source, and the positions of listeners. The reverberation must be adapted to the room's use. When the voice is central, a short reverberation time is preferred so that words remain intelligible. If the reverberation is too long, actors need to slow their rate of

FIGURE 1. THE CLASSICAL GREEK AMPHITHEATER of Tindari in Sicily, Italy. Renowned for its acoustics, it allowed vocal sounds to be heard to the last row, 60 meters from the stage. (Photo by Chris Lloyd, 2019, CC-BY-2.0.)

speech to remain understandable.[1] Whereas an ordinary living room may reverberate for a fraction of a second, a concert hall's reverberation time is typically around two seconds, and a cathedral's can exceed six seconds.

At the end of the 19th century, American physicist Wallace Clement Sabine laid the foundations of architectural acoustics by establishing a formula for calculating the reverberation time based on a room's volume and the acoustic properties of materials present. Today's architectural projects are developed using computer-aided design software, which allows engineers and acousticians to model the projects in two and three dimensions—sometimes including animations to provide virtual explorations of a space. Starting with architectural documents that detail the geometric characteristics of the performance hall and assumptions about the acoustic properties of its building materials, acousticians use those models to carry out predictive studies of the sound qualities of the future hall. The studies help them anticipate possible defects and propose modifications to architects. The same kinds of studies are used to understand the past acoustics of historical sites.

Modeling

Scientists have used physical and digital reconstruction methods for decades. But it's only recently that computational technologies have improved the quality and resolution of acoustic modeling sufficiently for researchers to tackle large-scale and complicated spaces. Sound in properly simulated spaces can be perceptually comparable to actual, on-site recordings.[2] Once created, the models can be modified to test acoustic conditions under different architectural configurations, source and listener positions, and use contexts. Acoustic simulations can be a powerful tool for historical

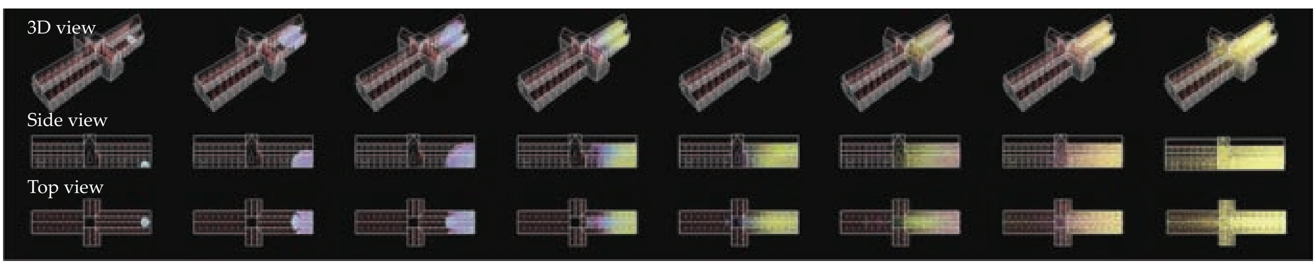


FIGURE 2. SIXTEENTH-CENTURY SOUNDS, RESTORED. A geometric-acoustic numerical simulation shows a graphical representation of the buildup and propagation of sound from an impulse-like source at the west end of the now ruined St Mary's Abbey Church in Museum Gardens, York, UK. Starting from the left, this sequence of images represents the first 400 ms of sound diffusing through the abbey. It takes roughly 6 s for the sound energy to decay to a level at which it would no longer be heard. That long reverberation time makes sounds heard far from the source much less intelligible than the dominant, direct sound of voices closer to the source. To hear music as it would have sounded in the abbey, listen to the soundcloud link at http://www.york.ac.uk/research/themes/sixteenth-century-acoustics. (Courtesy of Damian Murphy.)

FIGURE 3. AUDIO-VIDEO RENDERING over a head-mounted display. The portable system offers a fully immersive experience for historic reconstructions using either real-time, game-like explorations, or high-quality 360-degree video playback. (Adapted from B. Horan et al., *BMC Psychol.* **8**, 61, 2020.)

studies; they provide researchers with a sensory presentation of sound that had only been available earlier through descriptions.[3,4]

The transparent nature of acoustics is ideal for studying the layered nature of history in architectural sites. A geometrically accurate 3D model that incorporates the acoustic properties of relevant construction materials allows engineers to predict how the acoustics of a space will evolve as its geometry or materials change over time. In fact, changes that occur through the introduction, modification, or removal of material because of decay, renovation, or natural catastrophe can be incorporated into the model from documented evidence. Acousticians also examine changes in how a site is used in the context of the society's culture and customs over time.

Most existing approaches to numerical simulation for acoustic-heritage studies use one or more geometric-acoustic modeling techniques.[5] In geometric acoustics, sound is assumed to travel in straight lines, similar to a ray of light, and to propagate along paths calculated from its interaction with the three-dimensional model geometry of the environment; see, for example, the numerical simulation in figure 2. The result is a close approximation to the acoustic response of the modeled environment for a given set of conditions. However, results at low frequencies—typically below 500 Hz—are often less accurate than at high frequencies, as geometric-acoustic methods are less able to model the wave-like behavior of sound.

The wave-behavior limitation is an area of active research; an alternative approach is to use a numerical method to directly solve the underlying equations of wave motion.[6] Although

EUROPEAN EXEMPLARS: ACOUSTIC CULTURAL HERITAGE

Over the past couple of decades, the techniques of archaeological acoustics have become prevalent in historic research and in exploring the lost acoustic environments of significant but now damaged or destroyed buildings or performance venues. We outline a few recent and ongoing projects in Europe.

In *Re-sounding Falkland* (https://resoundingfalkland.com/), artists David Chapman and Louise K. Wilson collaborated with the Falkland estate in Scotland in 2010 to explore how sound can be used to understand and interpret the history of existing landscapes. The most significant challenge in the project was to create a three-dimensional model and an auralization of the Temple of Decision, a now-ruined structure on a hill overlooking the estate. Little is known about that 19th-century folly, and the acoustic reconstruction was informed by what the artists discerned from the ruins that remain, the fragments of documented evidence that could be found, and what is known about the construction of similar buildings.[4]

In France, the *ECHO* project, spanning the topics of voice, acoustics, and theatrical listening, examined the acoustical evolution of several important theater sites and was a tool between 2013 and 2018 for historians to test hypotheses.[3] Virtual reconstructions of the acoustics at Abbey St Germain-des-Prés and Notre Dame cathedral were carried out, with Notre Dame simulations available as virtual concert "fly-throughs," for public demonstrations. See the image above and the video link at http://www.lam.jussieu.fr/Projets/GhostOrchestra.

Bretez is an interdisciplinary project that explores the 3D setting, with audio and visual historically inspired reconstructions of the 18th-century Paris soundscape. Based on historical archives, maps, and other sources, it aims to construct an authentic multisensory immersive environment. See the YouTube video, https://youtu.be/YP__1eHeyo4. Other recent works have used both physical and numerical reconstructions, such as with the prehistoric Lascaux cave.[14] The experimental virtual archaeological acoustics (EVAA) project is developing a real-time dynamic simulator for musicians to experience the acoustics of historic performance spaces (see the image on the title page of this article).

In Italy in 1996, the Gran Teatro La Fenice in Venice burned to the ground. Two months beforehand, acoustical measurements had been made of the opera house,[15] work that set the stage for a reconstruction project intended to preserve the theater's original acoustical properties. A few years later, the Waves IR1 project captured the acoustical fingerprints of more than 100 theaters, churches, and caverns all around the world with the aim of preserving their unique acoustical behavior for posterity.[16]

The *ERATO* project is another milestone for archaeoacoustic research. Its goal was to analyze and compare the acoustical properties of ancient Greek and Roman theaters.[17] The *SIPARIO* project, at the University of Parma, aims to create real-time acoustical renderings of historical theaters for performers, who can then sing or play an instrument in a virtual environment that re-creates the theater's visual and acoustic presence.

FIGURE 4. AN EQUIRECTANGULAR IMAGE taken from a simulation of a performance at the Théâtre de l'Athénée in Paris. An audience is represented visually by animated mannequins and aurally by audio samples of spatially distinct intermittent coughing, which provide a sense of an occupied theater. Such 360° images can be rendered in a head-mounted display or shown on an immersive projection screen. (Adapted from the ECHO project: http://www.lam.jussieu.fr/Projets/CloudTheatre; rendering by David Poirier-Quinot.)

more accurate, such methods are too expensive computationally to offer a complete solution. Computer programs can take hours or weeks to reach final results across the full audio bandwidth for a large complex space. Hence, hybrid methods that combine geometric-acoustic, wave-based, and other statistical approaches are also an area of current research.[7]

Virtual reality

Auralization is the sound equivalent of visualization. The auditory presentation of an acoustical numerical model, through auralization over headphones or speaker arrays, lets users experience a site's acoustic properties as if they were actually there.[8]

The acoustics of a space is immersive and—due to the nature of auditory perception—egocentric, or individual, in contrast to the visual perception of an object, which can be viewed from outside. Today's technologies for creating an acoustical space use ideas and methods from virtual-reality (VR) systems, and are often integrated with visual rendering, as images have been shown to affect auditory perception.[9] Two approaches have emerged: One uses dedicated rooms equipped with large loudspeaker arrays and projection screens surrounding the listener, and the other uses VR helmets or head-mounted displays (HMDs) like the one worn in figure 3.

An HMD is equipped with binaural headphones and a device that tracks the position and orientation of the wearer's head. With that head-tracking functionality, it can often achieve a more stable reproduced soundscape. But for both loudspeakers and HMDs, the first level of realism requires processing three degrees of freedom (DOF)—the movement of the listener's head around three axes. A higher level of realism can be obtained with six DOF, which accounts for the wider movement of a listener around the virtual space.

To accommodate simple, three-DOF rendering, the acoustic characteristics—either measured or modelled—are typically represented as a higher-order Ambisonic (HOA) multichannel stream. Ambisonics is a hierarchical, spatial audio format that decomposes the sound field into spherical harmonic signals, which are then decoded to the listener's speaker setup or directly to headphones, with optional head tracking.[10] (For some background on the technique, read the book review in PHYSICS TODAY, June 2020, page 52.) That decomposition is an efficient way to represent the spatial distribution of sound at a fixed point. The higher the order, the greater the number of spherical harmonics, and the more accurate the spatial information.

Today, the balance between realistic reproduction and computational complexity is usually found using third-order Ambisonics, which requires 16 audio channels. For three-DOF video rendering, a simple panoramic camera will do the job, capturing photos or videos in the so-called equirectangular format, shown in figure 4. That type of image can also be easily generated by a computer in the case of purely virtual rendering.

The HOA stream is then "decoded" for either loudspeaker array or headphone reproduction. Similarly, equirectangular images must also be processed for either projector screens surrounding the listener or an HMD

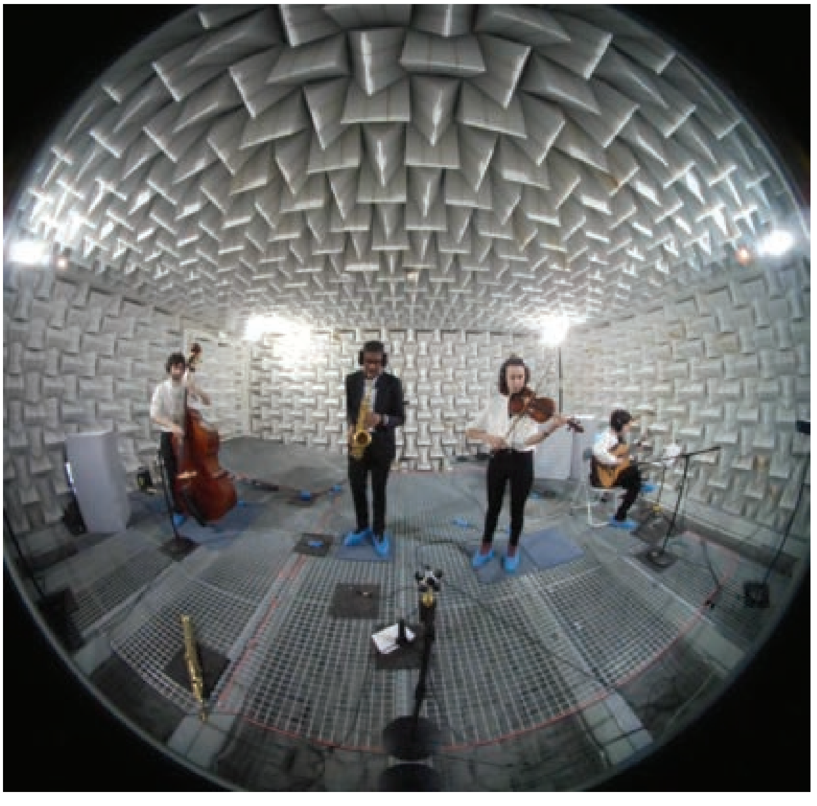


FIGURE 5. AN ANECHOIC CHAMBER recording session of a quartet. (Adapted from www.lam.jussieu.fr/Projets/AVAD-VR; photo by David Thery.)

screen. The result is a virtual space in which the listener can freely look around in every direction, having the impression of actually being inside the scene. Precise temporal and spatial matching between visual rendering and acoustic rendering is crucial for ensuring consistency between the senses and to avoid nausea from VR sickness.

HMDs use head-tracked binaural rendering of the HOA audio stream, which allows the VR system to take advantage of individualized filtering functions that represent the acoustical response of each listener's head and ears. Those individual filters are called head-related transfer functions, and they can be measured in special labs or numerically simulated from geometric data. The use of such filters in tandem with low-latency head-tracking devices provides such a realistic experience that many listeners cannot distinguish reproduced and real sounds.

Six-DOF systems are still experimental, particularly for rendering existing acoustical spaces. They require capturing the sound simultaneously with numerous microphones scattered around the area where virtual listeners can move. Although some laboratories are now attempting that approach, it's mostly used with computer-simulated acoustical renderings,[11] in which software simultaneously computes the sound field at hundreds of different listening points.

For loudspeaker rendering, most systems offer only two DOF for movements—or five DOF in total—allowing listeners to move freely in the room but keeping their heads at the same elevation above the floor. That arrangement is used in the "Museum of Reproduced Sound" in Parma, Italy. Sala Bianca, a room in the museum, has an array of 189 loudspeakers hidden in the walls. In the case of rendering over HMDs, the latest generation of devices can reliably provide six-DOF tracking of position and orientation of a listener's head. Most applications use software, such as Unity and Unreal Engine, originally made for video games.

Rooms do not sound on their own; they require a sound source. And presentation can have a significant effect on the listener's experience. The choice of appropriate source material—intermittent coughs from a virtual theater audience, say, or footsteps in a hall—helps put the site in its cultural and societal context; it may also connect the site to its surroundings. But sounds used in reconstructions should be recorded "dry," with no surrounding acoustic environment. The use of an anechoic room, like that shown in figure 5, achieves that objective by capturing only the direct sound, which can then be injected into the virtual reconstruction.

Taking into account the natural behavior of sources, such as the movements of actors or musicians on stage, improves the realism of reconstructions.[12] Modern systems are no longer limited to static acoustic sources. Ones are readily available that can capture or render sound sources rotating around three axes. And the next generation of systems will increasingly allow for six DOF, in which the sound sources can also move freely in space.

Reflections on historical reconstructions

Exploring cultural heritage through acoustic digital reconstruction provides historians, musicologists, and others with a perspective not available using more established research methods. Furthermore, it brings a powerful means of communicating and delivering memorable, meaningful, and most importantly, multisensory experiences. The effectiveness of digital reconstruction is evident through the range of projects undertaken throughout Europe—see the box on page 35.

Generally, auralization is only one particular, static representation of how an environment sounds. It's a snapshot in time, and the final result depends on the limitations of the recording systems and techniques as much as on the design criteria applied to the project. In the development of a model for any heritage space, the auralization is only as good as the research documenting its history.

Perhaps most importantly, our perception of a particular auralization reflects our own contemporary culture and our own prior experience of sound events. As with many historical conceptualizations, the final results are both created from and perceived through our modern state of mind.

Funding has been provided by the European Union's Joint Programming Initiative on Cultural Heritage project PHE, The Past Has Ears.[13]

REFERENCES

1. J. W. Black, *J. Acoust. Soc. Am.* **22**, 174 (1950).
2. B. N. J. Postma, B. F. G. Katz, *Virtual Real.* **19**, 161 (2015); B. N. J. Postma, B. F. G. Katz, *J. Acoust. Soc. Am.* **140**, 4326 (2016).
3. B. N. J. Postma, S. Dubouilh, B. F. G. Katz, *J. Acoust. Soc. Am.* **145**, 2810 (2019); B. F. G. Katz, M.-M. Mervant-Roux, *Rev. Sci./Lett.* **6** (2019), doi:10.4000/rsl.1645.
4. D. Murphy et al., *Internet Archaeol.* **44** (2017), doi:10.11141/ia.44.12.
5. L. Savioja, U. P. Svensson, *J. Acoust. Soc. Am.* **138**, 708 (2015).
6. J. van Mourik, D. Murphy, *IEEE/ACM Trans. Audio, Speech, Lang. Process.* **22**, 2003 (2014); S. Bilbao et al., *Comput. Music J.* **43**, 15 (2019).
7. A. Southern, D. T. Murphy, L. Savioja, *J. Acoust. Soc. Am.* **145**, 2770 (2019); A. Southern et al., *IEEE Trans. Audio, Speech, Lang. Process.* **21**, 1940 (2013).
8. S. Harriet, D. T. Murphy, *Acta Acust. United Acust.* **101**, 798 (2015).
9. D. Thery et al., in *Virtual Reality and Augmented Reality: 14th EuroVR International Conference, EuroVR 2017, Laval, France, December 12–14, 2017, Proceedings*, J. Barbic et al., eds., Springer (2017), p. 105; B. N. J. Postma, B. F. G. Katz, *Proc. Mtgs. Acoust.* **30,** 015008 (2017).
10. J. Daniel, S. Moreau, "Further study of sound field coding with higher order Ambisonics," paper presented at the Audio Engineering Society 116th Convention, 8–11 May 2004.
11. B. N. J. Postma et al., "Virtual reality performance auralization in a calibrated model of Notre-Dame Cathedral," paper presented at EuroRegio2016, 13–15 June 2016; N. Mariette et al., "SoundDelta: A study of audio augmented reality using WiFi-distributed Ambisonic cell rendering," paper presented at the Audio Engineering Society 128th Convention, 22–25 May 2010.
12. B. N. J. Postma, H. Demontis, B. F. G. Katz, *Acta Acust. United Acust.* **103**, 181 (2017).
13. B. F. G. Katz, D. Murphy, A. Farina, in *International Conference on Augmented Reality, Virtual Reality and Computer Graphics: 7th International Conference AVR 2020, Lecce, Italy, September 7–10 2020, Proceedings, Part II*, L. De Paolis, P. Bourdot, eds., Springer (2020), p. 91. See also "Welcome to 'The Past Has Ears,'" http://pasthasears.eu.
14. D. E. Commins, Y. Coppens, T. Hidaka, *J. Acoust. Soc. Am.* **148**, 918 (2020).
15. L. Tronchin, A. Farina, *J. Aud. Eng. Soc.* **45**, 1051 (1997).
16. A. Farina, R. Ayalon, in *24th AES International Conference on Multichannel Audio, the New Reality*, Audio Engineering Society (2003), art. no. 38.
17. A. Farnetani, N. Prodi, R. Pompoli, in *ERATO Project Symposium: Audio Visual Conservation of the Architectural Spaces in Virtual Environment*, S. Erdoğan, ed., Yildiz Technical University (2006), p. 27. PT

ULTRASOUND IN AIR

Experimental studies of the underlying physics are difficult when the only sensors reporting contemporaneous data are human beings.

Timothy G. Leighton

DONNA PADIAN

Timothy Leighton is the professor of ultrasonics and underwater acoustics in the Institute of Sound and Vibration Research at the University of Southampton in the UK.

Airborne ultrasound is becoming more prevalent in public places. Some individuals are complaining of adverse effects, including nausea, dizziness, tinnitus, fatigue, migraines and persistent headaches, and an uncomfortable feeling of "pressure in the ears."[1] Reduced technological costs have led to ultrasound being incorporated into new technologies beyond the pest deterrents that have been used for decades[2] (see figure 1). But tracking the increased prevalence of ultrasound in public spaces is difficult because there are no requirements to report it. A complicating factor is that the symptoms individuals attribute to ultrasound can be caused by other means. Whether someone has been exposed to ultrasound and to what level and for how long is often unclear, which makes the causal relationship difficult to establish.

When researchers use humans in experiments, verifying the adverse effects caused by ultrasound is inherently problematic. Susceptibility to ultrasound varies between subjects, and ethical research standards constrain human testing to exposures at lower levels and for shorter durations than one might experience from, say, a commercial pest deterrent. The interpretation and repeatability of experiments is complicated by patterns of strong scattering and diffraction around instruments, stands, and ears. Such difficulties contribute to insufficient data for guidelines: Only one, defined and labeled as interim in 1984, relates to public exposures, and it used data from adults to establish a guideline that should also protect children.[3]

The physical effects of such radiation, which in air has a wavelength of around 1 cm and scatters readily off skin, are therefore not well understood. That uncertainty has resulted in underinformed occupational guidelines and standards regulating ultrasound exposure.

Inherent problems in human experimentation

One challenge in obtaining and interpreting repeatable experiments is illustrated in figure 2, which shows the calculated real part of a pressure wave scattered off an ear. The wave originated at a point source placed 1 m from the opening of the ear canal and in the same horizontal plane. At low frequencies (200 Hz and 2 kHz in the figure), small changes in the relative positions of the ear and the source have little effect on the real part of the pressure that reaches the ear canal's entrance. That enables the listener to localize the source based on the signals detected at both ears.

However, at higher frequencies, the wavelengths—about 2 cm at 18 kHz and 1.5 cm at 23 kHz—become comparable with the length scales of the ear canal itself and the folds in the pinna. Slight movements of the ear relative to the source cause large variations in the received signal through scattering, diffraction, and ear canal resonances. The listener consequently struggles to locate and describe the signals, and even when the same individual is used in controlled experiments, results are difficult to replicate. Although researchers might attempt to make stands and rooms anechoic, or free from echo, scattering and diffraction at the head, pinna, and ear canal remain. Different humans have individually shaped pinnae, which makes reproducing results even more difficult. That discrepancy is further compounded by huge variation in individuals' middle- and inner-ear sensitivities to ultrasound.

People tend to lose high-frequency sensitivity as they age, but the vast variation in sensitivity to ultrasonic frequencies is poorly appreciated. Recent data suggest 5% of people ages 40–49 have hearing thresholds at 20 kHz that are at least 20 dB more sensitive than the median of those in the 30–39 age bracket.[4] That means the quietest sound some listeners in the older age bracket can hear at 20 kHz has 1/100 the power of the quietest sound that is audible to the average listener in the younger group. Moreover, 5% of those ages 5–19 are reported to have a 20 kHz threshold that is 60 dB lower than the median for those in the 30–39 age group (see figure 3), meaning the quietest sound they can hear has 10^{-6} as much power as the quietest audible sound for the average person in the 30–39 age bracket.

If the propensity for adverse effects does in fact increase with hearing sensitivity—which is so far unproven, given the ethical issues with experimentation—then that factor of 1 million casts into doubt any process that bases exposure guidelines for everyone, including children, on the propensity of most adults to experience adverse effects.

Figure 4 illustrates the exposure of young people. In the sketches, a minority of the students can point to the source of a sound that they reported to be unpleasant and distracting. About half the class cannot hear the sound, nor can any of the school staff. Those students that could hear it did not all agree on its location. Fortunately, the teacher took the matter seriously and, after a web search for an adviser, contacted me. I showed the teacher how to equip smartphones with an app that could detect the sound.[1,5] A maintenance team arrived and, with equipped phones, located the source—a defective motion sensor for the classroom lights that was supposed to be functioning at 40 kHz—and removed it. (See reference 5 for the full story.)

Many of those who have complained to me of adverse effects expressed frustration that their complaints had been dismissed by those who either could not detect the sounds or questioned whether such ultrasound, if it existed, could affect humans. Skeptics have raised physics-based arguments,[6] including that the intensity of ultrasound in air would be too low to cause physical effects and that more than 99% of ultrasound is reflected by skin. But both arguments are equally true of audio-frequency sound, which is why humans have evolved a complex hearing and balance system—one that could conceivably respond to ultrasound.

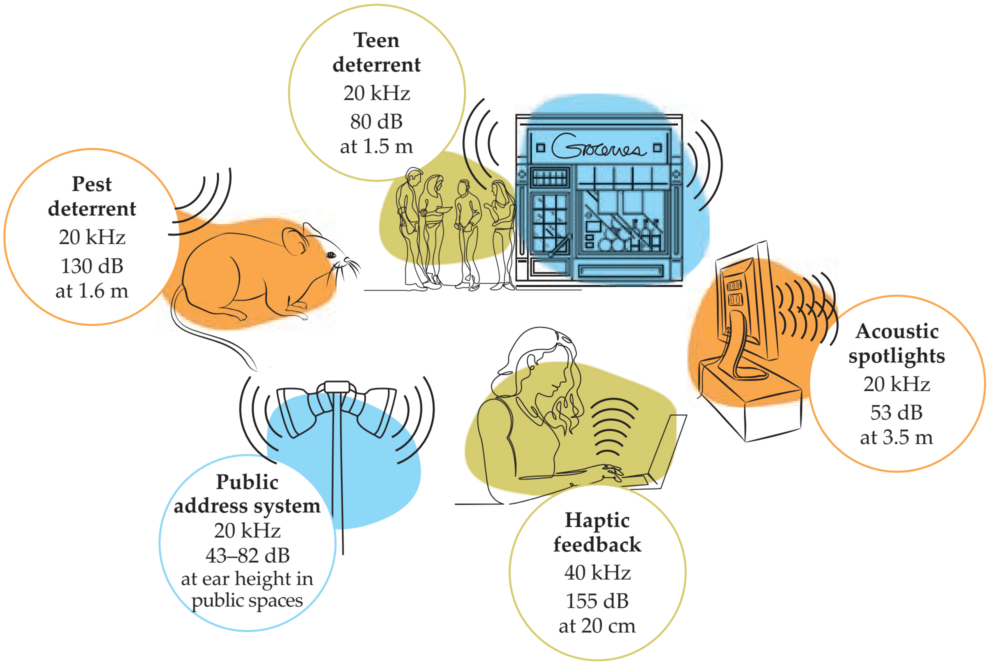


FIGURE 1. EXAMPLES OF INCIDENTAL AND DELIBERATE HIGH-FREQUENCY EXPOSURE from commercial devices. Sound pressure levels (SPLs) referenced to 20 μPa are shown. Pest deterrents use ultrasound to scare away birds, rodents, and insects. Similarly, teen deterrents are used to discourage young people from congregating by exploiting their sensitivity to high-frequency sounds. Many public address systems that use speakers to alert people to, say, a fire or a bomb threat emit a 20 kHz tone to aid in monitoring the system's function. Acoustic spotlights deliver targeted sound through two overlapping high-intensity ultrasonic beams whose interference produces a low-power audible sound. Devices with haptic feedback use modulated ultrasonic beams to produce vibrating sensations. The sources shown can also emit other frequencies.[1] Comparison with figure 3 shows that some of the measured SPLs could be perceptible to certain listeners.[12,14–16,18] Although it is unknown whether ultrasound must be audible to produce an adverse effect, it is known that audible ultrasound can.[13] (Figure by Donna Padian, based on table 1 in ref. 5.)

Setting guidelines

The regulation of ultrasound raises the question, How is ultrasound defined? Establishing it as sound above the frequency that an individual can hear is unworkable, since the highest frequency varies between individuals. Although a person's in-air hearing sensitivity decreases sharply with increasing frequency in the low ultrasonic range, pure-tone thresholds of 88 dB referenced to 20 μPa have been recorded[7] at 24 kHz. The 20–24 kHz frequency range has been a popular choice for commercial devices that emit airborne ultrasound, possibly because the frequency is low enough to avoid excessive attenuation and still be effective for applications such as pest deterrence. Use of that frequency range also provides a misguided sense of safety, since conventional wisdom says that humans cannot hear above 20 kHz.

The International Commission on Non-Ionizing Radiation Protection's charter states that the organization provides guidance on protection from exposure to "acoustic fields with frequencies above 20 kHz (ultrasound)." Thus lower frequencies are left to other bodies to provide direction. However, all national and international bodies set maximum permissible levels (MPLs) for frequency ranges in a third-octave band (TOB)—a frequency band containing one-third of an octave and referred to by its center frequency.[1] The TOB centered on 20 kHz runs from 17.8 kHz to 22.4 kHz, so any MPL set to limit exposure at 20 kHz would equally apply throughout the band. That choice effectively sets the lower bound of the ultrasonic range[6,8] at 17.8 kHz, thereby including the 18 kHz tone that disturbed the students depicted in figure 4.

Those defending the placement of commercial ultrasound-emitting devices in public places often cite occupational MPLs. They claim that regulatory bodies have reached consensus around an MPL of 110 dB re 20 μPa for the frequencies of most interest above the 20 kHz TOB, namely 22.4–56.2 kHz. But it is not a consensus based on independent data sets; rather, it is the

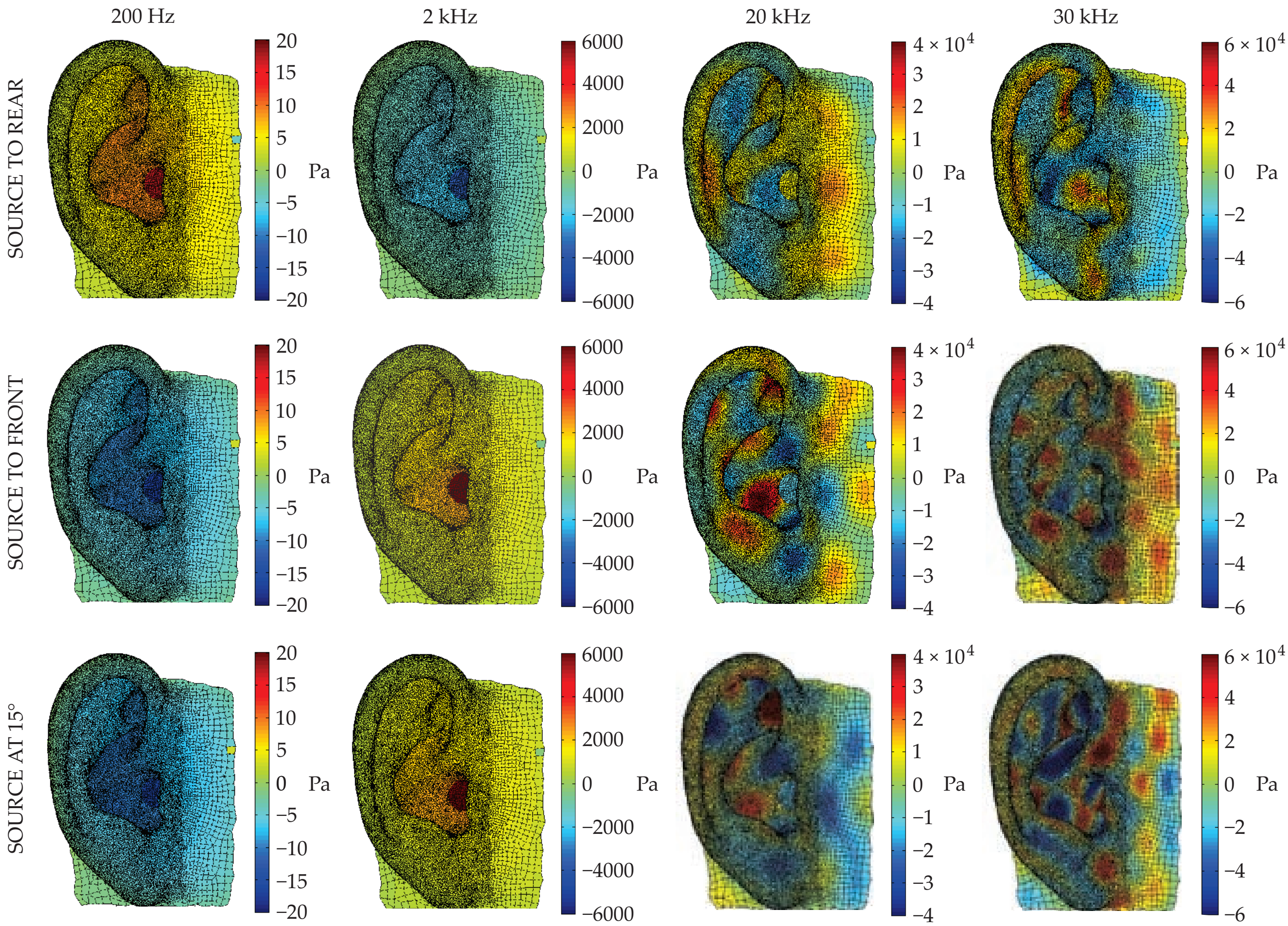


FIGURE 2. PRESSURE WAVES SCATTERING OFF AN EAR. The real part of the scattered pressure, in Pa, is calculated based on a point source placed 1 m from the ear canal's opening and in the same horizontal plane. At 200 Hz and 2 kHz, the source's location can be discerned as behind the listener (top row), in front of the listener (middle row), or angled 15° from the front position (bottom row). But at 20 kHz and 30 kHz, the wavelengths (around 1.7 cm and 1.1 cm, respectively) are of a similar size to the ear canal and pinna structures, and the signal varies greatly in response to small changes in the relative position of the head and source. Those two issues make experiments difficult to repeat and exposures tricky to reproduce. (Modeling by Erika Quaranta. Adapted from ref. 1.)

result of possibly under-resourced regulating bodies basing new guidelines on older ones.[1] The original data set underlying the guidelines appears to have been based on adult men who had worked in noisy environments. It was not large or diverse enough to typify those most in need of protection from public exposure—the minority who are adversely affected. They cannot be treated as outliers.

For the general public, such guidelines are clearly inappropriate. Millions of people might pass through or work in a railway station each day. If even a minority of those people are sensitive to ultrasound, they could number in the tens of thousands.[1] Their ages, exposure times and frequencies, medical histories, and possible health deterioration cannot be followed, and hearing protection cannot be enforced. MPLs should therefore be more conservative for public exposure than for occupational exposure.

In 2004 the US Occupational Safety and Health Administration (OSHA) reduced the appropriateness of their own guidelines[9] by voting to adopt two specific measures recommended by manufacturers through the American Conference of Governmental Industrial Hygienists. In the first, the ACGIH recommended setting MPLs by considering only one adverse effect: hearing loss caused by the subharmonics of ultrasonic frequencies. But inducing hearing loss likely requires a higher sound pressure level (SPL) compared with other adverse effects, including those experienced by the students shown in figure 4. The guideline also excludes hearing loss caused by most of the incident energy because subharmonics are usually weak relative to the fundamental frequency. OSHA nevertheless adopted the recommendation, stating, "These recommended limits (set at the middle frequencies of the one-third octave bands from 10 kHz to 50 kHz) are designed to prevent possible hearing loss caused by the subharmonics of the set frequencies, rather than the ultrasonic sound itself."

Air versus water

The second OSHA resolution adopted directly from the ACGIH's recommendation draws on a physics argument. The national and international guideline on MPLs for occupational exposure above the 20 kHz TOB cluster around 110 dB re 20 μPa. OSHA's guideline states that the allowable limits could be increased by 30 dB, from 110 dB to 140 dB re 20 μPa, "when there is no possibility that the ultrasound can couple with the body by touching water or some other medium."

The reasoning behind that 30 dB increase was not given, but further inquiry identified the following argument: If a plane acoustic wave in air (density ρ_1 and sound speed c_1) were normally incident on water (density ρ_2 and sound speed c_2), then the proportion T of the incident intensity that is transmitted can be calculated[10] using the formula

$$T = 1 - (\rho_2 c_2 - \rho_1 c_1)^2 / (\rho_2 c_2 + \rho_1 c_1)^2.$$

Substituting the properties of air ($\rho_1 = 1.225$ kg/m^3, $c_1 = 343$ m/s) and water ($\rho_2 = 1000$ kg/m^3, $c_2 = 1500$ m/s), which is commonly used as a first-order model for soft tissue, gives $T \approx 0.001$. That is to say, only 1/1000 of the incident intensity in air is transmitted into the soft tissue, which equates to a 30 dB attenuation.

OSHA's second resolution only makes sense, however, if one assumes that the data informing the 110 dB re 20 μPa so-called consensus limit were taken with the transducer pressed against the subject's head or with both the head and the transducer immersed in water. They were not: Both the transducer and the head were in air. By applying the 30 dB allowance in 2004, OSHA gave the US the most lenient MPLs in the world. OSHA recently removed explicit mention of the 30 dB allowance from its webpage.[1]

A similar argument has been taken further by some manufacturers. In the early to mid 2010s, the company uBeam raised investment by advertising a system whereby ultrasound in air would wirelessly recharge mobile phones and other devices in conference venues, airports, hospitals, and other public spaces. It was coy regarding frequencies and intensities, but when pressed on safety, the company placed the following assurance on its website: "The power levels beamed are more than 50 times lower than the lowest ultrasound imaging exposure limits set by the FDA for medical imaging, making the system inherently safe and within all existing regulatory constraints."

Rather than considering the hazard from ultrasound in air, uBeam apparently relied on the Food and Drug Administration's guidelines for 1–30 MHz ultrasound in soft tissue, such as the womb. But that limit is inappropriate. The primary concerns in fetal imaging—namely, potential cavitation and heating—are not relevant for airborne ultrasound at the levels used in public places. In fact, the only paper the company cites for powering devices by ultrasound has no air in the propagation path.[11]

Manufacturers and academics must be transparent in their calculations, especially when using decibels to compare intensities in air and water, which often introduces two common errors.[2] First, the decibel is not an absolute measure, and the SPL uses different reference levels in air (20 μPa) and water (1 μPa). Second, the factors $\rho_1 c_1$ and $\rho_2 c_2$ that are implicit in converting SPLs to intensities differ between air and water by a factor of around 3500. Failure to account for those differences has led to erroneous conclusions, such as the suggestion "that the sound of the penis of the 2 mm-long freshwater insect *Micronecta scholtzi* rubbing against its abdomen 'reached 78.9 decibels, comparable to a passing freight train.' "[10]

Devices

Given the difficulties setting MPLs and measuring SPLs, it is hard to assess the safety of available ultrasound-producing devices. Data on SPL outputs are largely unavailable because manufacturers are not obligated to publish them. Even when measurements are made, standard procedures may be inadequate. Acoustic measurement standards require, for example, the use of an anechoic chamber that reduces acoustic reflections from the wall or mapping levels in a grid of 5 cm spacings.[8] To my knowledge, however, no chambers are certified as anechoic up to 30 kHz, and many ultrasonic sources have main beams too narrow to map[12] using spacings as large as 5 cm.

Scattering, the increased directionality of sound sources and detectors, and other complications with ultrasonic frequencies are not sufficiently accounted for when using protocols designed for lower frequencies. For decades, physicists and engineers have measured the outputs of devices such as pest deterrents. But when their measurements are cited, is it noted whether they used a class 1 sound level meter, the laboratory standard? And if a class 1 device was used, were the researchers aware

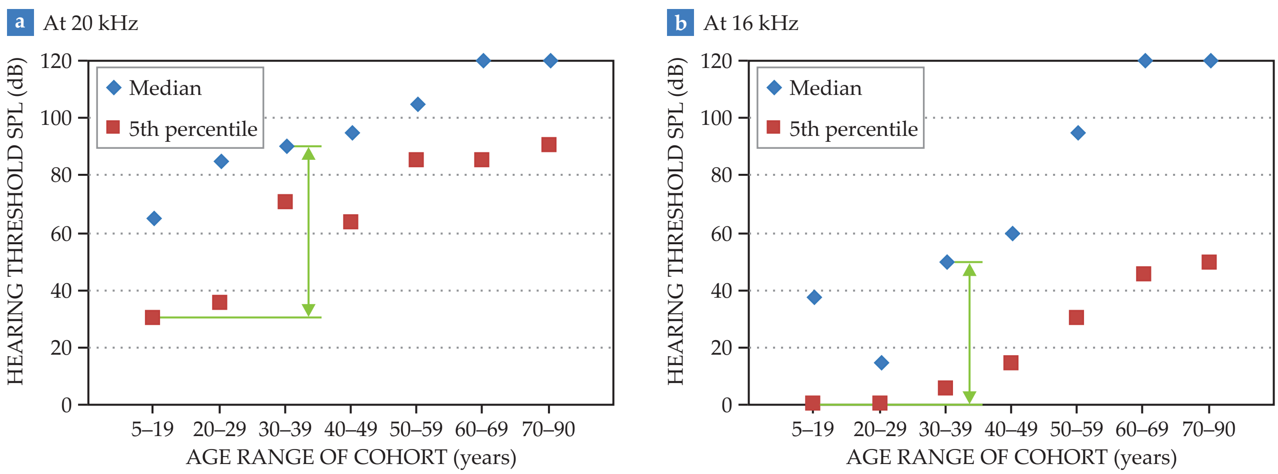


FIGURE 3. HEARING THRESHOLDS FOR PURE TONES. The median and fifth-percentile values based on 645 subjects are shown for **(a)** 20 kHz and **(b)** 16 kHz. The green arrows and lines indicate the difference between the median in the 30–39 age group and the fifth percentile in the 5–19 age group. Data at 0 dB and 120 dB are influenced by the instrumentation thresholds and saturation limits and so contribute less reliably to the statistics. All thresholds are sound pressure levels referenced to 20 μPa. (Data from ref. 4; figure adapted from ref. 1.)

FIGURE 4. LOCATING AN ULTRASOUND SOURCE IN A CLASSROOM. Some students in a classroom reported a persistent, high-pitched sound that was disrupting their work. But other students, teachers, and staff could not hear it. Sketches of photographs provided by the classroom teacher, Jill Zawatski, show the students who could detect the sound pointing to the location of the sound's source as they perceive it.

that although the device's standard performance acceptance limits are ±1 dB at 1 kHz, the limits in the 20 kHz TOB are +3 dB and −∞ dB? That means that for frequencies above 17.8 kHz, the device could underestimate the SPL to any degree without notifying the operator. (The levels in figure 1 were taken, where possible, from recent papers that used traceable calibrations.)

Human experimentation

In laboratory experiments, my colleagues and I found that some individuals had adverse effects to ultrasound they could hear.[13] We observed no adverse effects when individuals were exposed to ultrasound they could not hear *at the levels and durations used*. The italicized phrase is important because otherwise, the finding—one of the few in the field—could be misinterpreted as reaffirming the yet-unproven proposition that people cannot be affected by ultrasound they cannot hear.

In our tests, the SPLs and durations were restricted by ethics guidelines: The inaudible ultrasound produced only marginal results that could not be further evaluated by increasing exposures. Indeed, they did not even reach the SPLs and durations that humans might experience in some public places.[1,6,14,15]

It is ironic that for $20, you can place a pest deterrent in your garden and expose your neighbor's children to significantly higher ultrasound levels than we could use in controlled laboratory tests on carefully monitored adults. In a Tokyo restaurant, a pest deterrent's 20 kHz ultrasonic field reached 120 dB re 20 μPa directly under the source and 90 dB re 20 μPa some 15 m away. In a survey of volunteers, 31 out of 35 said they could hear it. Some had strong responses, including "my head may split" and "I will never come here again because of the pain in the ear."[14]

Continuing research efforts have confirmed the presence of ultrasound in public places;[15] measured the outputs of commercial sources;[14–18] documented the effects on humans;[13,14,17] and improved calibrations, standards, and procedures. Public attention increased in 2017 with claims of an ultrasonic attack on US and Canadian embassy staff in Cuba, although experts (myself included) remain skeptical that ultrasonic waves were the culprit.[8] The difficulties in proving or disproving the Cuban incident and the anecdotal claims of adverse effects from airborne ultrasound in public spaces clearly illustrate the importance of further study despite the challenges posed when humans are the only contemporaneous sensors.

I am grateful to Erika Quaranta for undertaking all the modeling in figure 2 showing the interaction of sound and ultrasound with the pinna.

REFERENCES

1. T. G. Leighton, *Proc. R. Soc. A* **472**, 20150624 (2016).
2. T. G. Leighton, *Prog. Biophys. Mol. Biol.* **93**, 3 (2007).
3. International Non-Ionizing Radiation Committee of the International Radiation Protection Association, *Health Phys.* **46**, 969 (1984).
4. A. Rodríguez Valiente et al., *Int. J. Audiol.* **53**, 531 (2014).
5. T. G. Leighton et al., *Acoust. Today* **16**(3), 17 (2020).
6. T. G. Leighton, *Proc. R. Soc. A* **473**, 20160828 (2017).
7. K. Ashihara et al., *Acoust. Sci. Technol.* **27**, 12 (2006).
8. T. G. Leighton, *J. Acoust. Soc. Am.* **144**, 2473 (2018).
9. Occupational Safety and Health Administration, *OSHA Technical Manual*, section III, chap. 5 (15 August 2013).
10. T. G. Leighton, *J. Acoust. Soc. Am.* **131**, 2539 (2012).
11. L. Radziemski, I. R. S. Makin, *Ultrasonics* **64**, 1 (2016).
12. M. Liebler et al., in *Proceedings of the 23rd International Congress on Acoustics*, M. Ochmann, M. Vorländer, J. Fels, eds., German Society for Acoustics (2019), p. 6338.
13. M. D. Fletcher et al., *J. Acoust. Soc. Am.* **144**, 2511 (2018); M. D. Fletcher et al., *J. Acoust. Soc. Am.* **144**, 2521 (2018).
14. M. Ueda, A. Ota, H. Takahashi, in *43rd International Congress and Exposition on Noise Control Engineering (Internoise 2014): Improving the World Through Noise Control*, J. Davy et al., eds., Australian Acoustical Society (2015), p. 6507; M. Ueda, A. Ota, H. Takahashi, in *Proceedings of the 11th Congress on Noise as a Public Health Problem, ICBEN* (2014), paper ID 4-16.
15. M. D. Fletcher et al., *J. Acoust. Soc. Am.* **144**, 2554 (2018); B. Paxton, J. Harvie-Clark, M. Albert, *J. Acoust. Soc. Am.* **144**, 2548 (2018); P. Mapp, *J. Acoust. Soc. Am.* **144**, 2539 (2018); F. Scholkmann, *Acoustics* **1**, 816 (2019).
16. C. N. Dolder et al., *J. Acoust. Soc. Am.* **144**, 2565 (2018); E. Conein to D. Howell, memorandum (12 April 2006), Gloucestershire Hospitals, ref. 06nstaffsp.doc, available at https://bit.ly/2XqFfCv.
17. A. van Wieringen, C. Glorieux, *J. Acoust. Soc. Am.* **144**, 2501 (2018).
18. C. N. Dolder et al., in ref. 12, p. 6359.

Ocean acoustics in the changing Arctic

Peter F. Worcester and Megan S. Ballard

Recent changes in ice cover and ocean stratification have been so large that acoustic measurements made during the Cold War no longer reflect current conditions.

MEGAN BALLARD

Peter Worcester is a research oceanographer emeritus at the Scripps Institution of Oceanography in San Diego, California. **Megan Ballard** is a research scientist at the University of Texas at Austin.

Interest in Arctic acoustics began in the early years of the Cold War when nuclear-powered submarines capable of operating for extended periods under the ice were first developed.[1] In 1958 the first operational nuclear submarine, the USS *Nautilus*, reached the North Pole. Shortly thereafter, military and academic scientists conducted Arctic acoustics research from ice islands and seasonal camps on the ice. The knowledge they gained was used to support submarine operations and develop antisubmarine warfare. Research at those ice camps continues today, but military interest in Arctic acoustics waned in the early 1990s at the end of the Cold War.

Researchers, however, are increasingly interested in the dramatic changes occurring in the Arctic Ocean in response to rising atmospheric concentrations of carbon dioxide and other greenhouse gases (see the article by Martin Jeffries, James Overland, and Don Perovich, PHYSICS TODAY, October 2013, page 35). Surface air temperature in the Arctic has warmed at more than twice the global rate over the past 50 years, and sea ice extent and thickness have declined dramatically.[2] Moreover, ocean stratification is changing as warmer waters encroach into the Arctic from the North Atlantic and North Pacific Oceans. Those changes have also affected acoustic propagation and ambient sound.

The rapidly changing Arctic Ocean

Since 1979, satellites with passive microwave sensors have measured the extent of sea ice—the area of the ocean that has at least 15% ice cover. Data from the US National Snow and Ice Data Center indicate that the monthly average sea-ice extent at its minimum in September declined from 7.67 million km^2 in 1984 to 4.32 million km^2 in 2019. (The lowest extent of 3.4 million km^2 of ice observed so far occurred on 16 September 2012.) Although sea ice varies considerably from year to year, a linear fit to the data from 1979–2019 yields a decline in the September monthly average of 12.9% per decade. The winter ice extent is also declining, albeit more slowly.

The presence of multiyear ice that survives the summer melt season is closely related to ice thickness and volume. Using a combination of satellite data and drifting buoys to derive the formation, movement, persistence, and disappearance of sea ice, researchers can estimate the age of multiyear ice. Figure 1 shows that the area covered by sea ice four years of age or older declined from 2.7 million km^2 in September 1984 to 53 000 km^2 in September 2019. The disappearance of multiyear ice has affected not only the average ice thickness but also pressure ridges atop the ice and ice keels that extend into the ocean, both of which form when large packs of floating ice floes collide. Old multiyear ice has larger pressure ridges than first-year ice, and the keels can extend tens of meters down. But those deep keels are disappearing with the multiyear ice.

The changes the Arctic is experiencing come in part from the warm, salty water of the Atlantic Ocean that enters the Arctic through the Barents Sea and the Fram Strait between Greenland and Spitsbergen, the largest island of the Svalbard archipelago.That water sinks and forms the relatively warm Atlantic layer approximately 150–900 m deep, as shown in figure 2. The Atlantic water makes its way throughout the Arctic Ocean: The bathymetry of the deep Arctic basin margins and ridges guides the water as it travels counterclockwise and deeper in the basin. Data from 1950 to 2010 show that since the 1970s the temperature of the Atlantic water has been steadily warming.[3]

In addition to the warming Atlantic layer, the western Arctic is influenced by waters entering from the Pacific Ocean through the shallow Bering Strait. Those waters comprise the relatively fresh Pacific Summer Water (PSW), between approximately 40 m and 100 m deep and having a local temperature and sound speed maximum, and the more saline Pacific Winter Water (PWW), characterized by a local temperature and sound speed minimum. Starting in 2000, the PSW warmed and thickened, and the depth of the PWW increased from 150 m to 200 m.[4]

Central Arctic acoustics

In general, sound speed increases monotonically with depth in the eastern central Arctic. As figure 2 shows, the sound speed profile is therefore upward refracting. As sound propagates, it interacts repeatedly with the ice. Some of the acoustic energy reflects with each interaction, some is scattered by the rough ice, and the rest is converted to compressional and shear waves within the ice. The resulting energy losses increase with frequency, which effectively makes the Arctic waveguide a low-pass filter.

During the Cold War, sound waves could only propagate to ranges in excess of a few hundred kilometers at frequencies below 30 Hz or wavelengths greater than 50 m because of losses due to interaction with the ice.[5] With the reduction in the amount of multiyear ice and the disappearance of the associated deep keels, higher-frequency sound can now propagate to greater ranges, but long-range propagation is still limited to relatively low frequencies. At very low frequencies, however,

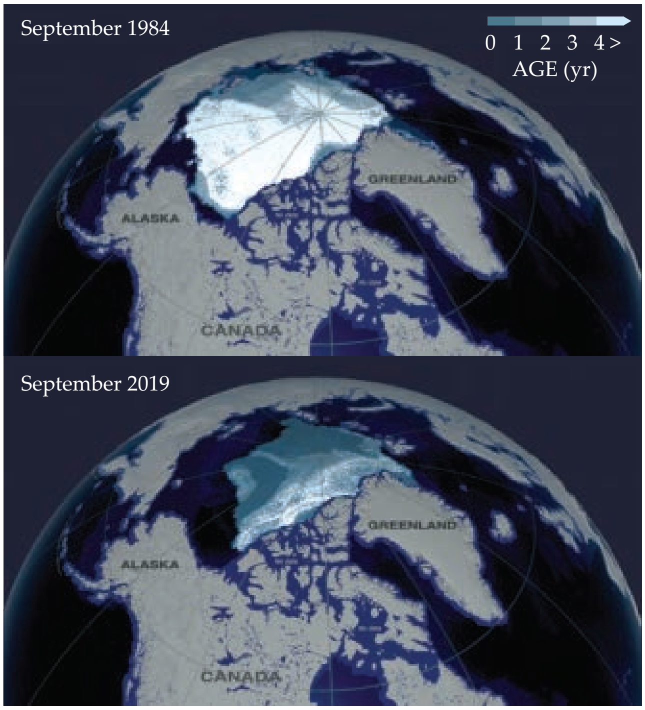


FIGURE 1. THE AGE OF ARCTIC SEA ICE in **(top)** September 1984 and **(bottom)** September 2019. The dark blue color represents first-year ice; the white, ice at least four years old. (NASA Scientific Visualization Studio, 2019.)

the sound begins to interact with the seafloor and lose energy to it. The resulting losses increase as frequency decreases, so only sound waves above a lower-frequency bound of 5–10 Hz can propagate over long distances.

Acoustic propagation in the western Arctic differs from that in the eastern Arctic because of the water from the Pacific. Sound speed increases with depth except for the sound speed minimum at the depth of the PWW, which forms an acoustic duct, visible in figure 2, that traps the sound energy and allows it to propagate over long distances without interacting with the ice cover or the seafloor. As the PSW has warmed and thickened in recent years, the sound-speed duct—sometimes referred to as the Beaufort Duct—has strengthened. An acoustic navigation and communication system deployed in the Beaufort Sea achieved ranges in excess of 400 km using 900 Hz sources deployed 100 m deep in the duct.[6] But the duct is sufficiently weak that signals at frequencies below a few hundred hertz are no longer fully confined to the duct and lose energy by interacting with the ice cover.

Various physical properties of the sea ice, including porosity, brine content, Poisson's ratio, Young's modulus, and the shear modulus, affect acoustic propagation. Those properties differ between first-year sea ice—the primary type in the modern Arctic—and multiyear sea ice. Compared with first-year ice, sound waves travel faster through multiyear ice primarily because of its lower brine content, which results from summer-melt processes. The shear-wave speed of first-year sea ice can be lower than the sound speed of the underlying seawater and consequently increase the transmission of sound out of the water column and into the sea-ice layer. Multiple studies have shown that the elastic properties of sea ice, particularly shear-wave attenuation, have the greatest influence on the surface reflection coefficient.[7]

However, the effects of the intrinsic properties of sea ice can be dominated by the effects of acoustic scattering because the under-ice surface has a rough topography.[8] Scattering of acoustic waves depends on the length scales present in the surface's roughness relative to the acoustic wavelength. Multiyear ice ridges can be significantly larger than first-year ice ones.[9]

Differences in roughness are possible even in the absence of ice ridges. The smooth underside of undeformed first-year ice differs from that of rugged multiyear ice, which is made by surface pools melting at uneven rates during previous summers. Increased roughness can cause additional coupling of acoustic energy from the ocean waveguide into the sea ice layer. The relative importance of the intrinsic reflection loss and the scattering loss depends on the acoustic frequency and the characteristics of the ice cover, which are variable in space and time.

Ambient sound

The sources and propagation of ambient sound in the Arctic differ substantially from those at lower latitudes. In temperate environments, low-frequency sound (20–500 Hz) is predominantly caused by distant shipping, and higher-frequency sound (500–100 000 Hz) is mostly from spray and bubbles associated with breaking waves. Ambient sound in the Arctic, however, is highly variable in space and time: It shows strong seasonal variations correlated to annual changes in ice cover. For example, anthropogenic sounds occur more frequently

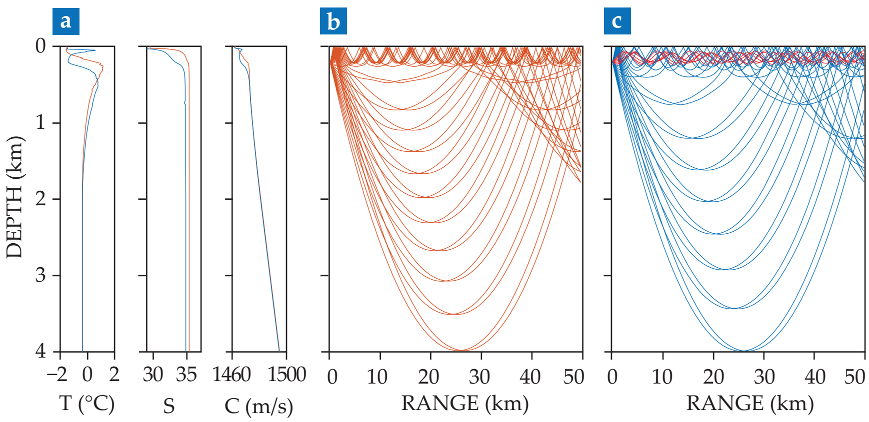


FIGURE 2. (a) THE TEMPERATURE (T), SALINITY (S), AND SOUND SPEED (C) PROFILES in the western (blue line) and eastern (orange line) Arctic Ocean were measured with Ice-Tethered Profiler instruments. Geometric ray paths for transmissions were generated from an acoustic source at 200 m depth in the **(b)** eastern and **(c)** western Arctic. Rays with positive (upward) and negative (downward) launch angles of the same magnitude give rise to the closely spaced ray pairs. In the western Arctic, a subset of the rays' paths that are contained in the Beaufort sound-speed duct are shown in red. (Data from R. Krishfield et al., *J. Atmos. Ocean. Tech.* **25**, 2091, 2008; J. M. Toole et al., *Oceanography* **24**, 126, 2011.)

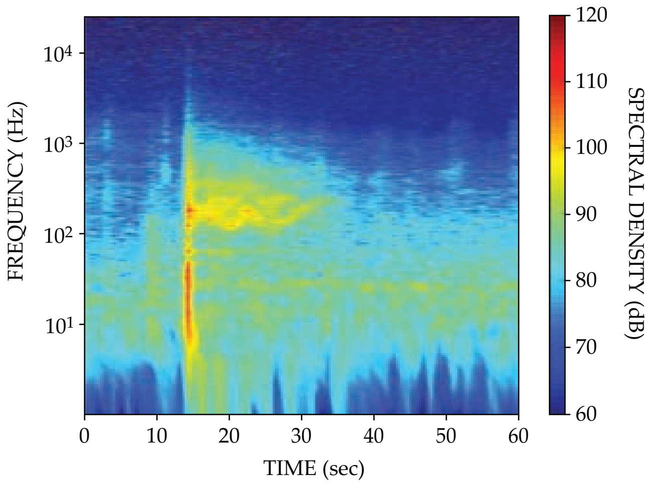


FIGURE 3. THIS TIME-FREQUENCY SIGNATURE of a local ice-cracking event approximately 400 m from the receiver was recorded on 16 March 2014 during the US Navy Arctic Submarine Laboratory Ice Experiment. The color bars are in dB relative to a reference level of 1 μPa^2/Hz. (Adapted from K. L. Williams et al., *IEEE J. Ocean. Eng.* **43**, 145, 2018.)

during the summer months when large portions of the Arctic Ocean are ice free. Additionally, observations of marine mammal vocalizations are correlated to species-specific seasonal migratory patterns, which often follow the ice edge.

During the winter and spring, the prevailing sound in ice-covered regions of the Arctic is largely generated when the ice cover deforms and fractures in response to wind, swell, currents, and thermal stresses.[9] Figure 3 shows the acoustic response of one such cracking event. Ice-fracturing processes can create some of the loudest underwater environments, but under calm wind conditions the Arctic can be one of the quietest places in the world's oceans. For example, under-ice ambient sound can reach levels 30 dB higher or 20 dB lower than the same ice-free summer location.[1] Ambient sound measurements have historically had a broad peak around 15–20 Hz, which results from the band-pass nature of sound propagation from distant sources.

Several studies have compared historical ambient sound measurements with recent data.[10] However, drawing conclusions from the comparisons is difficult given the high spatiotemporal variability of the Arctic sound field, the differing conditions under which the data sets were acquired, and the various metrics reported. Nevertheless, researchers expect the level and character of ambient sound in the Arctic to change as ice coverage and thickness decrease.

Changes in sound-generating mechanisms and acoustic propagation determine the evolving ambient sound field in the Arctic. Evidence already exists of an increase in anthropogenic activities from industries such as oil and gas exploration, fishing, shipping, tourism, and military activity. The once impassable Northwest Passage and the Northern Sea Route have recently seen increases in commercial ship traffic. Marine mammals are adapting to the changing conditions: Some species, such as Pacific Arctic beluga whales, are shifting their migratory patterns because of the longer ice-free season. Furthermore, the thinner and more mobile ice of the modern Arctic responds more readily to winds and is more easily broken up, which has increased wind-generated sound.

Acoustic remote sensing

Ocean acoustic tomography is a method that remotely senses the ocean interior by transmitting sound through it.[11] The speed at which sound travels depends on the ocean's temperature and velocity fields, and measurements of acoustic travel times therefore provide information on water temperature and current. Researchers use inverse methods on the measured travel times of the signals to infer the ocean's state. Travel times are inherently spatially integrating, which means they provide long-range horizontal and vertical averages and suppress the effects of small-scale oceanic variability that can contaminate point measurements. The application of long-range transmissions to measure ocean temperature is often referred to as acoustic thermometry.

During the Transarctic Acoustic Propagation (TAP) experiment in 1994, phase-modulated acoustic signals with a center

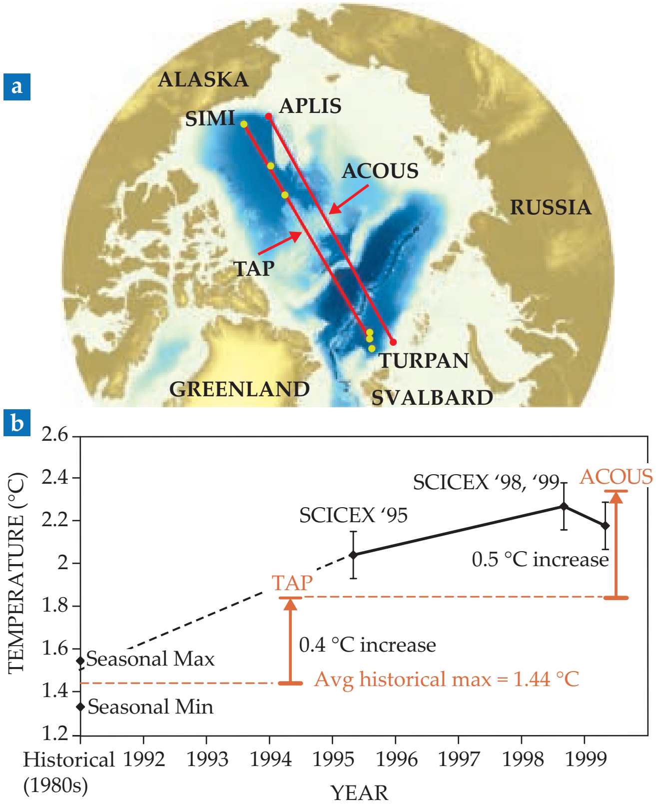


FIGURE 4. (a) THE MAP SHOWS THE ARCTIC OCEAN TRANSECTS of several acoustic remote-sensing initiatives, including the Transarctic Acoustic Propagation (TAP) experiment, the Arctic Climate Observations Using Underwater Sound (ACOUS) project, and the Coordinated Arctic Acoustic Thermometry Experiment (CAATEX, yellow dots). **(b)** The maximum temperature data of the Atlantic layer were obtained from historical climatology data, the 1994 TAP mode-2 acoustic travel times, the 1999 ACOUS mode-2 travel times, and the SCICEX 1995, 1998, and 1999 submarine transects. The uncertainties in the TAP and ACOUS temperatures are ± 0.2 °C and ± 0.25 °C, respectively. The change in maximum temperature (arrows) is inferred from the changes in acoustic travel times. (Adapted from B. D. Dushaw et al., in *Observing the Oceans in the 21st Century*, C. J. Koblinsky, N. R. Smith, eds., GODAE Project Office and Bureau of Meteorology, 2001, p. 391.)

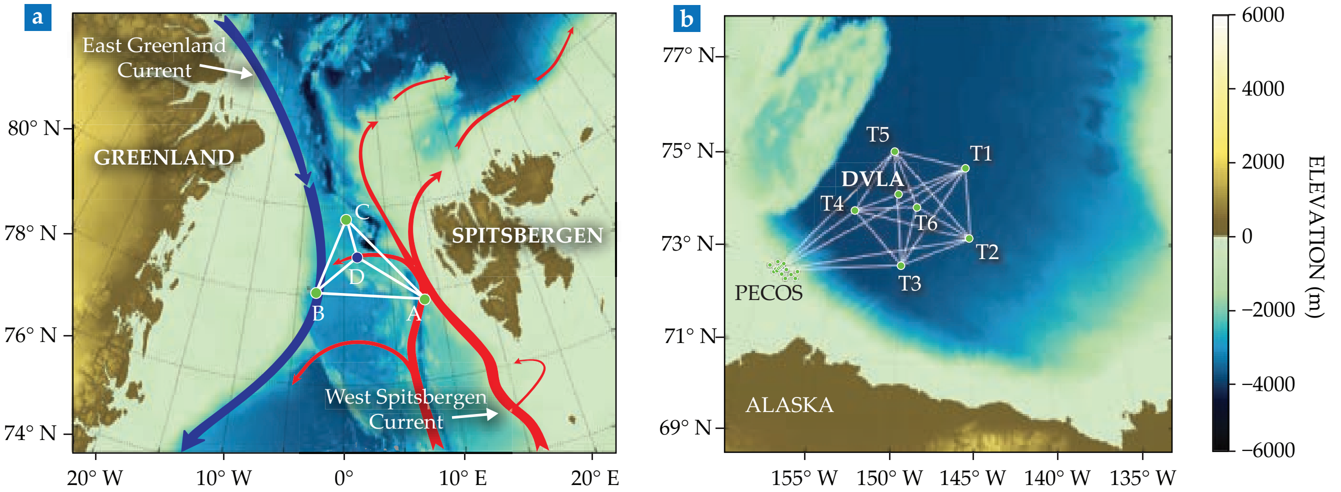


FIGURE 5. (a) THE ACOUSTIC TECHNOLOGY FOR OBSERVING THE INTERIOR OF THE ARCTIC OCEAN (ACOBAR) experiment ran from 2010 to 2012 to gather improved estimates of heat and freshwater transport in Fram Strait. Acoustic transceiver moorings (green) were located at A, B, and C. (Mooring C failed prematurely.) An acoustic receiver mooring (blue) was located at D. (Adapted from ref. 13.) **(b)** The geometry of the 2016–17 Canada Basin Acoustic Propagation Experiment included six acoustic transceiver moorings (T1 through T6) and a vertical line array receiver mooring (DVLA). The sound transmissions were also recorded by acoustic receivers located on the continental shelf and upper slope to the southwest of the array (PECOS). Some of the acoustic paths connecting the sources and receivers are shown in white. (Image by P. F. Worcester and M. S. Ballard.)

frequency of 19.6 Hz were transmitted by a source located at ice camp Turpan north of Svalbard. The signals traveled across the Arctic basin to receiving arrays at ice camps Narwhal in the Lincoln Sea and SIMI in the Beaufort Sea.[12] The distance from Turpan to SIMI was 2630 km, and the measured travel time of the acoustic normal mode that sampled the Atlantic layer was 2.3 ± 1.2 s shorter than that computed from long-term climate data. The result, summarized in figure 4, indicates a warming of 0.4 ± 0.2 °C across the basin. The warming was subsequently confirmed by measurements made from submarines during various efforts by the Science Ice Exercises Program and from icebreakers.

In October 1998, as part of the joint US–Russia Arctic Climate Observations Using Underwater Sound (ACOUS) project, another ultralow-frequency source (20.5 Hz) was moored in Franz Victoria Strait.[12] The transmissions were recorded on a moored receiver in the Lincoln Sea and by a receiver at ice camp APLIS in the Chukchi Sea in April 1999. The acoustic path to APLIS from the moored source was similar to that from Turpan to SIMI during the TAP experiment. ACOUS reported travel times 2.7 ± 1.3 s shorter than the TAP measurements: Scientists concluded that the Atlantic layer further warmed by 0.5 ± 0.25 °C.

Twenty years later, the 2019–20 US–Norway Coordinated Arctic Acoustic Thermometry Experiment (CAATEX) was conducted to repeat the basin-scale measurements made during the TAP and ACOUS experiments, as shown in figure 4a. Six moorings were installed, providing both basin-scale and shorter-range measurements. One mooring in the eastern Arctic and another in the western are acoustic transceivers, each composed of a 35 Hz acoustic source and a vertical receiving array. The remaining four moorings have only vertical receiving arrays.

Regional-scale tomographic experiments have been conducted in Fram Strait since 2008 and are shown in figure 5a. With a width of nearly 400 km and a sill depth of about 2600 m, Fram Strait is the only deep-water connection between the Arctic and the rest of the world's oceans. The West Spitsbergen Current transports relatively warm and salty Atlantic water into the Arctic to form the Atlantic layer. The East Greenland Current transports sea ice and relatively cold and fresh polar water out of the Arctic. Significant recirculation of the Atlantic water and intense variability in the center of the strait make it difficult to accurately measure ocean transport through the strait.

The primary goal of the experiments in Fram Strait is to determine whether acoustic measurements, when combined with other data and ocean models, provide improved transport estimates of the inflowing and outflowing water masses. All the experiments employed acoustic sources that transmitted linear, frequency-modulated signals with bandwidths of 100 Hz and center frequencies of approximately 250 Hz. The oceanographic conditions in Fram Strait produce complex acoustic arrival patterns.[13] Nonetheless, researchers succeeded in obtaining robust estimates of range and depth-averaged temperature along the acoustic paths using the measured travel times.[14] As a first step toward using the travel times to constrain ocean models, researchers have interpreted the structure and variability of the measured acoustic arrivals by comparing them to arrivals computed using sound-speed fields obtained from a high-resolution regional ocean model.[15]

The Canada Basin has experienced some of the greatest decreases in ice cover and the most significant changes in ocean stratification in the Arctic.[16] Six acoustic transceiver moorings similar to those used in Fram Strait and a vertical line array receiver mooring were deployed for the 2016–17 Canada Basin Acoustic Propagation Experiment (CANAPE), shown in figure 5b. The transmissions were also recorded by receivers on the continental shelf north of Alaska[17] and by acoustic Seagliders, autonomous underwater vehicles.

The goals of CANAPE were to understand how the changes in the ice and ocean stratification affected acoustic propagation and ambient sound. In addition, the data from the tomographic array provide information on the spatial and temporal variability of the upper ocean throughout the annual cycle and make it possible to assess whether acoustic methods, with other data and ocean modeling, can yield improved estimates of the time-evolving ocean state. The amplitudes of the receptions decrease in wintertime, likely a consequence of the increasing ice thickness, which reached a maximum of about 1.5 m in late winter, and of the changes in the structure of the Beaufort Duct. Whereas the transmission loss is highly variable over the year, the travel-time variability was roughly an order of magnitude smaller than is typical in midlatitudes at similar ranges. Compared with other oceans, the Arctic is a remarkably stable acoustic environment.

Multipurpose acoustic systems

Monitoring and understanding the rapid changes underway in the Arctic Ocean are of crucial importance for researchers assessing its role in climate variability and change. As the Arctic converts from largely perennial to seasonal ice cover, oil and gas exploration, fisheries, mineral extraction, shipping, and tourism will increase the pressure on the vulnerable environment. To inform and enable sustainable development and protect that fragile environment, scientists will need improved ocean, ice, and atmosphere data.

Gliders, profiling floats, and autonomous underwater vehicles that measure ocean temperature, salinity, and other variables cannot surface in ice-covered regions to use the Global Navigation Satellite System and relay data back to shore. Furthermore, measurements of sea-surface height using satellite altimeters, which provide important constraints on the ocean circulation at lower latitudes, cannot be obtained when the ocean is covered with ice.

Multipurpose acoustic systems, however, can operate beneath the ice. Such systems provide acoustic remote sensing of temperatures via ocean acoustic tomography, underwater navigation, and passive acoustic monitoring of natural and anthropogenic sounds. Such systems have a special role in making measurements of the rapidly changing Arctic Ocean and complementing and supporting other *in situ* observations.[18]

This article is based on the paper "Ocean acoustics in the rapidly changing Arctic" by one of us (Worcester), Matthew Dzieciuch, and Hanne Sagen, Acoustics Today, *volume 16 (2020), page 55.*

REFERENCES

1. D. Hutt, in *Advances in Ocean Acoustics: Third International Conference on Ocean Acoustics,* AIP Conf. Proc. Vol. 1495, AIP (2012), p. 56.
2. J. Stroeve, D. Notz, *Environ. Res. Lett.* **13**, 103001 (2018).
3. I. V. Polyakov, A. V. Pnyushkov, L. A. Timokhov, *J. Climate* **25**, 8362 (2012).
4. M.-L. Timmermans et al., *J. Geophys. Res. Oceans* **119**, 7523 (2014).
5. P. N. Mikhalevsky, "Acoustics, Arctic," in *Encyclopedia of Ocean Sciences,* J. H. Steele et al., eds., Elsevier (2001), p. 92.
6. L. Freitag et al., in *Proceedings of the OCEANS Conference and Exposition*, IEEE (2015), p. 913.
7. K. LePage, H. Schmidt, *J. Acoust. Soc. Am.* **96**, 1783 (1994).
8. G. Hope et al., *J. Acoust. Soc. Am.* **142**, 1619 (2017).
9. P. Wadhams, N. Toberg, *Polar Sci.* **6**, 71 (2012).
10. E. Ozanich et al., *J. Acoust. Soc. Am.* **142**, 1997 (2017).
11. W. H. Munk, P. F. Worcester, C. Wunsch, *Ocean Acoustic Tomography*, Cambridge U. Press (1995).
12. P. N. Mikhalevsky, A. N. Gavrilov, *Polar Res.* **20**, 185 (2001).
13. H. Sagen et al., *J. Acoust. Soc. Am.* **141**, 2055 (2017).
14. H. Sagen et al., *J. Geophys. Res. Oceans* **121**, 4601 (2016).
15. F. Geyer et al., *J. Acoust. Soc. Am.* **147**, 1042 (2020).
16. F. McLaughlin et al., *Oceanography* **24**, 146 (2011).
17. M. S. Ballard et al., *J. Acoust. Soc. Am.* **148**, 1663 (2020).
18. P. N. Mikhalevsky et al., *Arctic* **68**, 11 (2015).

PT

PRECISION MEASUREMENT GRANTS

The National Institute of Standards and Technology (NIST) anticipates awarding two new Precision Measurement Grants that would start on 1 October 2021, contingent on the availability of funding. Each award would be up to $50,000 per year with a performance period of up to three years. The awards will support research in the field of fundamental measurement or the determination of fundamental physical constants. The official Notice of Funding Opportunity, which includes the eligibility requirements, will be posted at **www.Grants.gov**.

Application deadline is tentatively February 2021. For details/unofficial updates see: physics.nist.gov/pmg.

For further information contact:

Dr. Joseph N. Tan, Ph.D., Manager
NIST Precision Measurement Grants Program
100 Bureau Drive, Stop 8422
Gaithersburg, MD 20899-8422, U.S.A.
email address: joseph.tan@nist.gov

NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

Max Planck Institute for the Physics of Complex Systems

mpipks

The Max Planck Institute for the Physics of Complex Systems in Dresden announces the opening of one or several

Postdoctoral Positions

in the area of condensed matter theory, to work with Roderich Moessner, Ashley Cook, Markus Heyl, David Luitz and Inti Sondemann. The areas of research range from strongly correlated Fermions and Bosons in and out of equilibrium, gauge theories, frustrated systems and topological/fractionalized phases of matter, via computational many-body physics, to quantum computation and machine learning.

The Institute provides a stimulating environment due to an active in-house workshop program and a broad range of other research activities. Strong experimental groups are nearby, in particular in the neighbouring Max Planck Institute for Chemical Physics of Solids.

To apply for a position, please fill the online application form (http://www.pks.mpg.de/CMpd21) and upload your application package (cover letter, curriculum vitae, list of publications, statement of research interests and research proposal as well as the three most relevant publications) in one pdf file. Please arrange for at least two letters of reference to be sent by **January 20, 2021** preferably to be submitted in pdf format online (http://www.pks.mpg.de/reference/); or by email to visitors@pks.mpg.de with subject line **CMpd21**; or by regular mail:

Max Planck Institute for the Physics of Complex Systems, Visitors Program, Nöthnitzer Str. 38, 01187 Dresden, Germany.

The Max Planck Institute aims to increase the number of women in scientific positions. Female candidates are therefore particularly encouraged to apply.

In case of equal qualifications, candidates with disabilities will take precedence.

MAX PLANCK GESELLSCHAFT

Colores Salsa (2018), Valerie Vescovi, www.valvescoviart.com

Michael Blaß, **Jost Leonhardt Fischer**, and **Niko Plath** are researchers at the University of Hamburg's Institute of Systematic Musicology in Hamburg, Germany.

Computational PHONOGRAM ARCHIVING

Michael Blaß,
Jost Leonhardt Fischer,
and Niko Plath

A general approach to analyzing audio files makes it possible to re-create the sound of ancient instruments, identify cross-cultural musical properties, and more.

Music archives store vast amounts of audio data and serve different interests. Some of the most prominent archives belong to the commercial streaming services that deliver music on demand to worldwide audiences. Users of Pandora, SoundCloud, Spotify, and other streamers access music through curated or user-generated playlists. Alternatively, they can search for an artist, song, or genre and scroll through the results on their devices.

The streamers' collaborative filtering algorithms also generate recommendations based on the preferences of other users. Motivated by commercial gain, major music companies promote their content by inducing the streaming services to suggest it. Their songs duly appear in heavy rotation. Combined, the filtering and the manipulation lead to so-called echo chambers, in which the recommendations in a distinct group of people gain a self-reinforcing momentum. Members end up listening to an identical subset of the music catalog.

Ethnomusicological archives are another type of audio repository. They aim to provide access to a wide variety of field recordings from different cultures around the world. The Smithsonian Folkways Recordings, the Berlin Phonogram Archive, and the Ethnographic Sound Recordings Archive (ESRA, esra.fbkultur.uni-hamburg.de), in which the three of us are involved, are just three examples. The reasons for maintaining such archives are various. Some are dedicated to preserving musical cultural heritage; others focus on education.

Repositories of a third type store the sounds of musical instruments. The institutions that maintain them focus on providing access to large sets of well-recorded sound samples. One use of such data is for the physical modeling of musical instruments. Researchers solve the differential equations of continuum dynamics to create realistic sounds that mimic the recordings. To make that possible, archives store the responses of instrument parts to impulsive forces and imaging data—such as CT scans of instruments or computer-aided design (CAD) models of them. Also stored are textual data from bibliographical research, such as instrument descriptions, provenance, and technical drawings. The databases are valuable for investigations of historic instruments whose surviving examples are unplayable.

The increasing demand in digital humanities for big data sets has led to another type of archive as cultural institutions rush to digitize their musical recordings and other holdings. Unfortunately, when such projects are complete, the result is often a vast trove of uncurated data.

The subject of our article, computational phonogram archiving, has emerged as a unified solution to tackling some of the generic problems of music archives, such as classifying recordings, while also addressing their particular shortcomings, such as echo-chamber playlists.[1] Its primary approach is to analyze, organize, and eventually understand music by comparing large sets of musical pieces in an automated manner.

As we describe below, the first step in computational phonogram archiving involves using extraction algorithms to transform music into numerical representations of melody, timbre, rhythm, form, and other properties. In the second step, machine learning algorithms derive mood, genre, meaning, and other higher-level representations of music.

From audio files to sonic content

Music archives consist of audio files and accompanying metadata in text format. Metadata typically include extra-musical information, such as the artist's name, genre, title, year of recording, and publisher. Archives of ethnographic field recordings also collect and store metadata about the origin of the audio. That information could include geographic region, GPS-derived latitude and longitude, the ethnic identity of the performers, and additional notes about the circumstances of the recording. If an archive focuses on preserving old recordings, metadata might also include information about the condition of the original media.

By using a search engine, users can query an archive's database through its metadata. But when the metadata are incomplete or absent, an alternative strategy comes into play.[2] Algorithms applied to musical recordings compute numerical representations of the characteristics of the audio data. At the most basic level, the representations, called audio features, quantify specific signal properties, such as the number of times per second the waveform crosses the zero-level axis and the amount of energy in various frequency bands.

To compute audio features that correlate with the auditory perception of human listeners, researchers employ methods drawn from the field of psychoacoustics.[3] Since Hermann von Helmholtz's groundbreaking derivation in 1863 of the Western major scale from perceptions of musical roughness, psychoacoustical investigations have explored the diverse relationships between physical properties of sound and the effects they have on auditory perception. The perception of loudness, for example, is not only proportional to amplitude, it is also strongly dependent on frequency and on whether or not a sound's component frequencies are integer multiples of each other.

Search and retrieval engines that interrogate audio files can directly compare data sets only when they are the same size. One way to circumvent that limitation is to resolve the natural time dependence of the audio features. To that end, several applications use central moments, such as the expected value and vari-

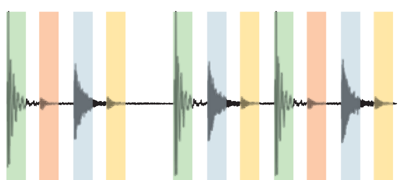
FINDING RHYTHM WITH A HIDDEN MARKOV MODEL

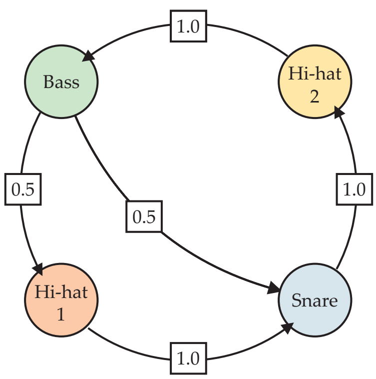
Researchers at the Ethnographic Sound Recordings Archive use a hidden Markov model (HMM) to derive a representation of a piece's rhythm in terms of the succession of timbres it includes. An onset detection algorithm estimates the time point of the beginning of each note. Feature extraction is performed only around those time points.

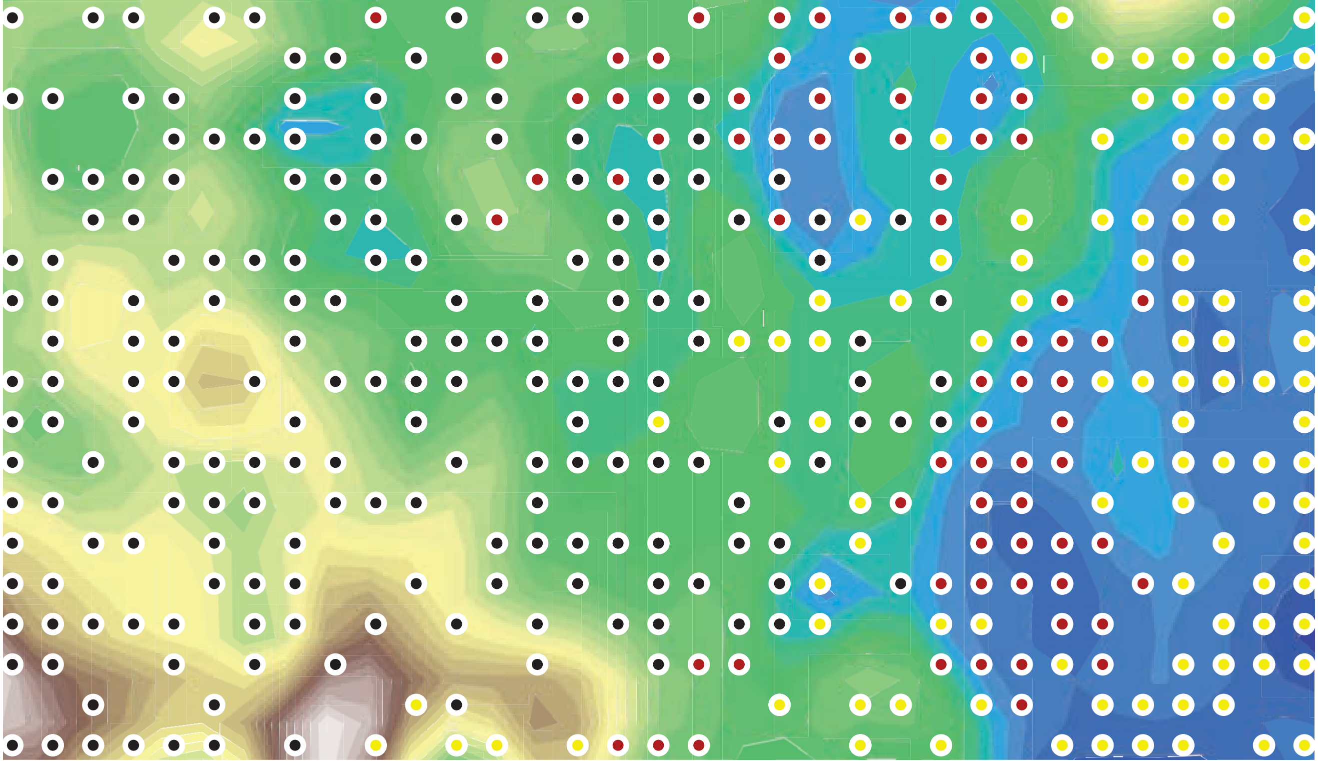
The training procedure considers the probabilities of changing from one timbre to another and then adapts the model's transition probability matrix. The matrix therefore embodies the internal structure of a trained rhythm. As such, it functions as a prototype of the overall "global" rhythm of a piece of music. The original time series is one instance of the global rhythmical profile.

How the process plays out is represented by the accompanying figures. The top one illustrates a drum groove as a waveform diagram. Colored rectangles mark the regions subject to feature extraction. The colors correspond to the three instruments that played the groove. Two of them, bass drum (green) and snare drum (blue), are treated individually. The third instrument, the hi-hat, is split into two: hi-hat 2 (yellow), which is the cymbal's pure sound, and hi-hat 1 (coral pink), which is a polyphonic timbre that combines the pure hi-hat sound with the decay of the bass drum.

The lower figure shows a graphical representation of the rhythmical profile. Each node corresponds to an HMM state. Edges represent the node transitions with probabilities greater than zero. The HMM quantifies a counterintuitive finding: Timbre influences how rhythm is perceived.







ance, to aggregate over the time dimension. In subsequent stages of the interrogation, those values replace the original features.

Another approach is to model the distribution of an audio feature by using a mixture model. In such a case, the estimated model parameters stand in for the original time series. A similar but more specialized approach is the use of a hidden Markov model (HMM).[4] The parameters of a trained HMM provide insight into the process that might have generated a particular instance at hand (see the box on page 52). Such a representation is especially useful in archives that feature music from different cultures, because its generality allows for an unbiased, extracultural viewpoint on the sound.

Machine learning and music similarity

Audio features consist of vectors—that is, sets of points—in a high-dimensional space. How can such a point cloud help users navigate through an archive's contents? The key idea is that similar feature vectors represent similar regions of the space in which they reside. The vectors' mutual distance is therefore a natural measure of their similarity. The same is also true for the musical pieces that the vectors characterize. If their feature vectors are close together, their content is similar. Methods from artificial intelligence enable the exploration of enormous sets of feature vectors.[5]

A straightforward approach, at least for retrieving records that resemble a particular input, is to use the *k*-nearest neighbors algorithm, which assigns an object to a class based on its similarity to nearby objects in a neighborhood of iteratively adjustable radius. Another way to classify the feature space is to partition it such that the members of a particular partition are similar. Such cluster analyses are conducted using methods like the *k*-means and expectation–management family of algorithms, agglomerative clustering, and the DBSCAN algorithm.

However, to truly explore the inherent similarities in the feature space, users need to somehow move through it. For that to be feasible, the complexity of the space has to be mitigated by reducing its dimensionality. Principal component analysis achieves that goal by projecting a data set onto a set of orthogonal vectors that account for the most variance in the data set.

FIGURE 1. A SELF-ORGANIZING MAP (SOM) trained with the timbre features of audio files in the Ethnographic Sound Recordings Archive (ESRA). Circles mark the position to which the SOM assigns the audio feature vectors of individual audio files. The circles' color denotes which of three ESRA collections the file belongs to. The clustering of the collections reflects the audio files' noise floor, which, in turn, is related to the recording technology. In particular, the oldest recordings (yellow dots) tend to be the noisiest (blue contours).

An alternative approach, the self-organizing map (SOM),[6] is popular for reducing dimensionality in applications that entail retrieving musical information.[7] The SOM estimates the similarity in the feature space by projecting it onto a regular, two-dimensional grid composed of artificial "neurons." Users of the archive can browse the resulting grid to explore the original feature space.

Another popular approach for reducing dimensionality is to use an embedding method, which maps discrete variables to a vector of continuous numbers. By coloring data points that correspond to the available metadata, users can cluster the embedding visually. Popular methods for doing this are *t*-distributed stochastic neighborhood embedding (*t*-SNE) and uniform manifold approximation and projection (UMAP).

Figure 1 illustrates an example of the above approach. ESRA currently holds three collections of recordings from different time spans: 1910–48, 1932, and 2005–18. When a SOM clustered the assets of ESRA by timbre, it revealed that timbre features appear to be sensitive to the noise floor introduced by different recording equipment—that is, the SOM segregates input data by the shape of the noise. Some pieces lack dates. One of the SOM's applications is to estimate the year a piece was recorded.

Physical modeling

Physical modeling of musical instruments is a standard method in systematic musicology.[8] The models themselves are typically based on two numerical methods: finite difference time domain (FDTD) and finite element method (FEM). Computed on parallel CPUs, GPUs (graphical processing units), or FPGAs (field-programmable gate arrays), the two methods can solve multiple coupled differential equations with complex boundary conditions in real time.

Solving the differential equations that embody an instrument's plates, membranes, strings, and moving air is an iterative process that aims to reproduce musical sound. The solutions can also deepen understanding of how instruments produce sound. Instrument builders use models to estimate how an instrument would sound with altered geometry or with materials of different properties without having to resort to physical prototypes.

Brazilian rosewood, cocobolo, and other traditional tonewoods have become scarce because of climate change and import restrictions. Another use of models is to identify substitute materials. Conversely, some historical instruments, notably organs and harpsichords, no longer sound as they used to because their wood has aged and their pipes, boards, frames, and other components have changed shape. Models can reproduce the instruments' pristine sound.

A machine learning model can estimate both the range of possible historical instrument sounds and the performance of new materials by learning the parameter space of the physical model. After inserting the synthesized sounds that result into an archive and applying the methods of computational sound archiving, the essential sound characteristics of instruments can be estimated. Although the field is still emerging, a scientifically robust estimation of historical instrument sounds is in sight for the first time.

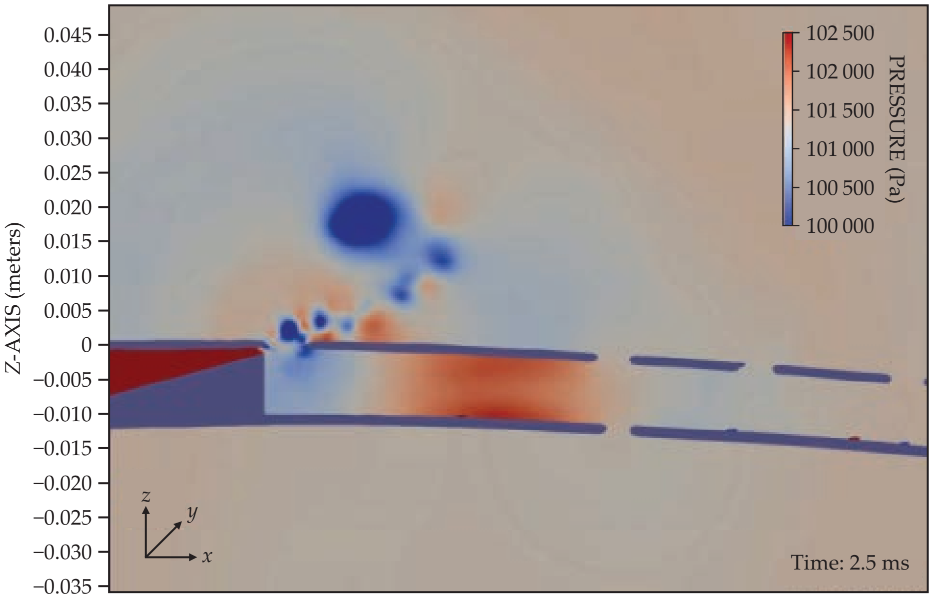


FIGURE 2. A VISUALIZATION OF THE PRESSURE FIELD around the mouthpiece of a simulated 15th-century bone flute 2.5 ms after the virtual player starts blowing.

Resurrection of a bone flute

As simulations proliferate, the need emerges to compare virtually created sounds with huge collections of sounds of real instruments. An example of such a comparison is the investigation of the tonal quality of an unplayable musical instrument: an incomplete 15th-century bone flute from the Swiss Alps.

Like its modern relative the recorder, a bone flute consists of a hollow cylinder with holes for different notes. The sound is produced when air is blown through a narrow windway over a sharp edge, the labium. The windway and mouthpiece are formed by a block that almost fills the flute's top opening. The bone flute in our example is unplayable because the block, which was presumably made of beeswax, has not survived.

The first step in resurrecting the flute's sounds is to scan the object using high-resolution x-ray CT scans. The 3D map characterizes both the object's shape and its density. The data are then ingested into the archive, where they are subjected to segmentation. Widely used in medical imaging to identify and delineate tissue, segmentation allows substructures with homogeneous properties to be extracted from a heterogeneous object.

Each voxel (a pixel in 3D space) is allocated to exactly one of the defined subvolumes. Since substructures are generated according to their homogeneous properties, the respective surface data adequately represent the body. What's more, replacing myriad voxels with a modest number of polygon meshes leads to a massive reduction of data.

A set of 3D meshes of instrument parts can be stored in the database for subsequent research tasks. The meshes are even small enough to be shared via email. The reconstructed polygon surface mesh is transformed into a parametric model and further modified using CAD. Once the virtual flute is made, it can be numerically fitted with a selection of differently shaped blocks that guide the air flow to strike the labium at varied angles.

And that's just what we did. Our model considered three angles (0°, 10°, and 20°) at which the blown air left the windway and struck the labium. It also considered two speeds (5 m/s and 10 m/s) at which the blown air entered the mouthpiece. We used the initial speed because the main dynamics that characterizes the generation and formation of tone—what players of wind instruments call articulation—takes place in the initial few milliseconds of the blowing process.

The sound optimization was carried out with OpenFOAM, an open-source toolbox for computational fluid dynamics and computational aeroacoustics. We set the tools to work solving the compressible Navier–Stokes equations and a suitable description of the turbulence. The six different combinations of angle and blowing speed were calculated in parallel on the 256 cores of the Hamburg University computing cluster. Time series of relevant physical properties were sampled from data and analyzed to find the optimal articulation of the instrument.

Figure 2 visualizes the pressure field in the generator region of the virtually intact bone flute. The simulated sound pressure-level spectra of the six combinations of angle and blowing speed revealed that one, 20° and 5 m/s, produced the best and most stable sound, a finding that was corroborated with measurements of a 3D-printed replica of the bone flute. The five other combinations led to attenuation of the transient and unstable or suboptimal sound production.

Virtually generated instrument geometries can be subjected to physical modeling, just like a real instrument. In particular, we could investigate more complex dynamics such as overblowing, a technique that Albert Ayler, John Coltrane, and other avant-garde saxophonists used to extend the instrument's sonic landscape.

Further applications

Computational phonogram archiving can have political implications. Some multiethnic states define ethnic groups according to their languages and how or whether the languages are related to each other. Like language, music is an essential feature of ethnic identity. Using the methods presented in this article, ethnologists can investigate ethnic relations and definitions from the perspective of sound in ways less susceptible to human prejudice than traditional methods are.

It's also possible with the tools of computational phonogram archiving to reconstruct historical migrations of music's rhythms, timbres, and melodies. Another application is the investigation into the ways that individual musical expressions relate to universal cross-cultural musical properties such as the octave.

The doctrine of the affections is a Baroque-era theory that sought to connect aspects of painting, music, and other arts to human emotions. The descending minor second interval, for example, has always been associated with a feeling of sadness. Other, perhaps previously unrecognized associations—historical and contemporary—could be revealed through computational phonogram archiving. Composers and compilers of movie soundtracks might profit from that approach. Given a model trained on features of musical meaning, they could choose appropriate film music by matching a particular set of emotions and signatures.

The analysis of the underlying psychoacoustic features could illuminate the musical needs of historical eras. Whether a piece of music had a particular function is often subject to debate, which could be resolved through automated comparison with the contents of a music archive.

Indeed, the power and promise of computational phonogram archiving derives from its ability to address problems on different scales. It enables researchers to compare large-scale entities, like musical cultures, and small-scale entities, like differences in the use of distortion by lead guitarists.

REFERENCES

1. R. Bader, ed., *Computational Phonogram Archiving*, Springer (2019).
2. I. Guyon et al., eds., *Feature Extraction: Foundations and Applications*, Springer (2006).
3. H. Fastl, E. Zwicker, *Psychoacoustics: Facts and Models*, 3rd ed., Springer (2007).
4. W. Zucchini, I. L. MacDonald, *Hidden Markov Models for Time Series: An Introduction Using R*, Chapman & Hall/CRC (2009).
5. J. Kacprzyk, W. Pedrycz, eds., *Springer Handbook of Computational Intelligence*, Springer (2015).
6. T. Kohonen, *Self-Organizing Maps*, 3rd ed., Springer (2001).
7. M. Leman, *Music and Schema Theory: Cognitive Foundations of Systematic Musicology*, Springer (1995).
8. R. Bader, ed., *Springer Handbook of Systematic Musicology*, Springer (2018).

PT

Massachusetts Institute of Technology

Faculty Position

The Physics Department at the Massachusetts Institute of Technology (MIT), located in Cambridge, Massachusetts, invites applications for a faculty position in **Astrophysics**, specializing in the study of exoplanets. Faculty members at MIT conduct research, teach undergraduate and graduate physics courses, and supervise graduate and undergraduate participation in research. Candidates must show promise in teaching as well as in research. A Ph.D. in physics or physics-related discipline is required by the start of employment. Preference will be given to applicants at the Assistant Professor level.

Current astrophysics faculty are active in broad areas of observational and theoretical astrophysics, including optical/IR, radio, and high energy astronomy; studies of exoplanets; gravitational wave science; and observational and theoretical cosmology. MIT hosts the Kavli Institute for Astrophysics and Space Research, whose faculty and research staff contribute instrumentation for and conduct research using several facilities including the Transiting Exoplanet Survey Satellite (TESS), the Laser Interferometer Gravitational-Wave Observatory (LIGO), the Magellan telescopes, the Hydrogen Epoch of Reionization Array (HERA), the Canadian Hydrogen Intensity Mapping Experiment (CHIME), the Chandra X-ray Observatory, and the Neutron Star Interior Composition Explorer (NICER), as well as an in-house high-performance computing cluster. We seek candidates who have both the potential for and strong commitment to innovation and leadership in teaching and mentoring undergraduate and graduate students, and share the Principles of our Community.

A complete application includes a cover letter, curriculum vitae, a 1- to 2-page statement on research and one on teaching and mentoring, and three letters of recommendation to be uploaded to https://academicjobsonline.org/ajo. Recognizing that educational experiences of all students are enhanced when the diversity of their backgrounds is acknowledged and valued, we ask candidates to articulate (in the teaching and mentoring statement, and, as appropriate, in the cover letter or research statement) their views on inclusivity and equity as they pertain to teaching, mentorship, research, and service. Enquiries should be directed to Prof. Claude Canizares, Search Committee Chair, crc@mit.edu.

MIT is an equal employment opportunity employer. All qualified applicants will receive consideration for employment and will not be discriminated against on the basis of race, color, sex, sexual orientation, gender identity, religion, disability, age, genetic information, veteran status, ancestry, or national or ethnic origin.

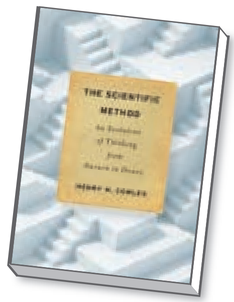
http://web.mit.edu

STEVE VIDLER/ALAMY STOCK PHOTO

Charles Darwin's study in his home, Down House, outside London.

The importance of hypothesis

The Scientific Method
An Evolution of Thinking from Darwin to Dewey
Henry M. Cowles
Harvard U. Press, 2020. $35.00



Some history of science studies are groundbreaking because they examine a previously neglected topic. Others are innovative in their use of new archival material. Still others use well-known sources to tell a new story about subjects previously tackled by scholars. Henry M. Cowles's *The Scientific Method: An Evolution of Thinking from Darwin to Dewey* falls into the last category. His novel account provides a big-picture look at the development of the scientific method across the natural, medical, and social sciences in both Britain and the US during the 19th century.

Cowles argues that Charles Darwin's evolutionary theory birthed an entirely new concept of the scientific method that shaped the thinking of biologists, anthropologists, neurologists, philosophers, psychologists, and educators in the late 19th century. He begins by stating flatly on the first page that "there is no such thing as the scientific method, and there never was." While that may seem to be an unpromising start for a book examining the history of the scientific method, Cowles's intention is to identify the history of the scientific method as a powerful myth.

According to *The Scientific Method,* a significant development in that history took place in the early 19th century, when philosophers and scientists such as John Stuart Mill, William Whewell, and John Herschel increasingly emphasized the role of hypothesis in the scientific method and rejected the notion that it could only lead to wild speculation. Darwin picked up on the importance of hypothesis, but he broke with the earlier figures by placing it within an evolutionary framework. For Darwin, Cowles argues, both species and theories evolved through a process of trial and error.

The testing of hypotheses was an integral part of the competition between scientific theories. Cowles shows how Darwin's theory of pangenesis, put forward in his *The Variation of Animals and Plants Under Domestication* (1868), was not an aberration. It was reminiscent of Darwin's unpublished early theorizing (1836–44) about evolution. Darwin believed that an integral part of the scientific method was formulating bold ideas that helped explain a variety of phenomena.

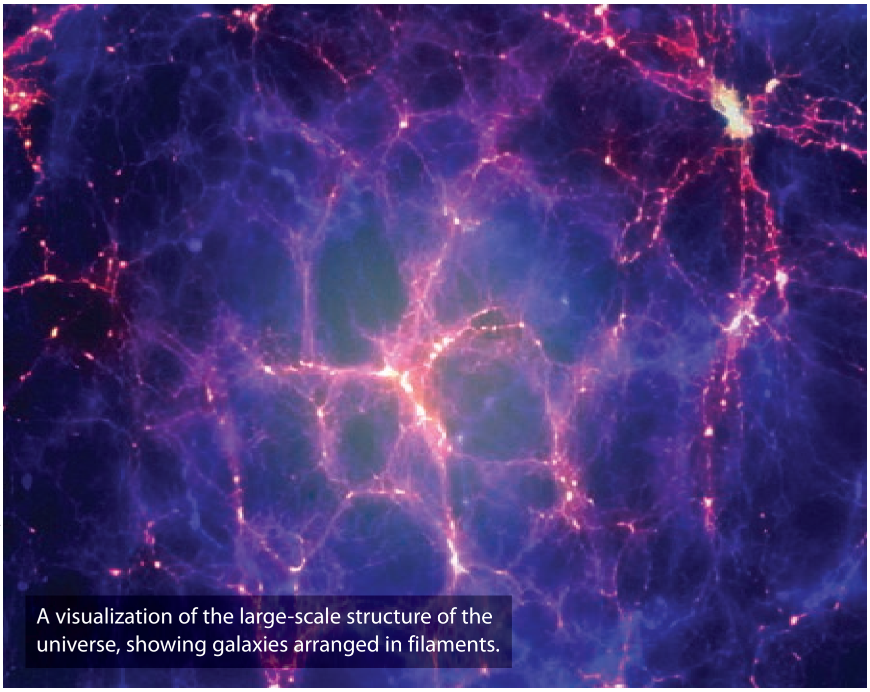
British figures from various disciplines subsequently embraced the notion that science and nature used the same method; they applied it explicitly to an understanding of how science evolved. For example, in his account of mental evolution, Herbert Spencer sought to construct a reflexive natural history of the mind that would include science as an object of study. Science itself became a tool used by the adaptive mind to cope with a constantly evolving world. Neurologists like William Benjamin Carpenter and John Hughlings Jackson and anthropologists like Edward Burnett Tylor explored the ramifications of the concept of the scientific method for their respective disciplines.

Halfway through the book, Cowles shifts his focus to the US to follow the development of Darwin's insight in the field of evolutionary psychology. He begins by depicting the meetings of the Metaphysical Club, an informal discussion group of philosophers that met in Cambridge, Massachusetts, during the 1870s, and uses them to discuss the importance of hypothesis and the nature of science to American thought. (Curiously, Cowles does not compare the Metaphysical Club to its British counterpart, the Metaphysical Society, even though such a comparison would serve as a way to transport his readers across the Atlantic.) The section examines philosophers and psychologists such as Charles Peirce, William James, Edward Thorndike, G. Stanley Hall, Ernest Lindley, and John Dewey.

As Cowles argues on page 185, studies of animal intelligence in comparative psychology "expanded outward to new minds and turned inward in new ways simultaneously, seeking the source of their scientific thinking in observable animal behaviors." Psychologists began to study children in the hope that they would reveal the developmental roots of human reasoning. After discussing John Dewey's interest in experimentalism, Cowles concludes by outlining how the school of psychology Dewey founded, functionalism, became the dominant line of thought in the US by the end of the 19th century.

The Scientific Method is an absorbing read that illuminates the history of the natural and social sciences in Britain and the US. It features nuanced readings of important scientific figures from a new perspective. Well-argued, accessible, and based on extensive research, Cowles's hypothesis about the transformation of the scientific method by evolutionary theory should win the struggle for existence in Darwin's "tangled bank" of scholarship on 19th-century science.

Bernard Lightman
York University
Toronto



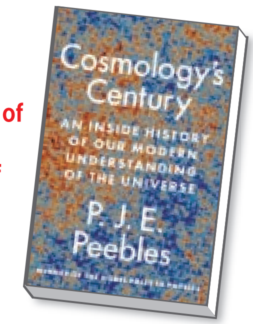
A visualization of the large-scale structure of the universe, showing galaxies arranged in filaments.

SCIENCE PHOTO LIBRARY/ALAMY STOCK PHOTO

Looking back on a modern scientific revolution

Cosmology's Century
An Inside History of Our Modern Understanding of the Universe

P. J. E. Peebles
Princeton U. Press, 2020. $35.00



The 20th century saw a revolution in our observational and theoretical understanding of the universe. P. J. E. Peebles's latest book, *Cosmology's Century: An Inside History of Our Modern Understanding of the Universe*, offers an inside look at how the pillars of modern cosmology were built. Using detailed historical research, quotes from key scientists, and his own recollection, the author presents an engaging story of the twists and turns taken on the long road toward our modern understanding of cosmology.

Peebles is uniquely positioned to recount that tale, as he is the key player in the story's genesis. After making monumental contributions to essentially every aspect of our current cosmological model, he received the 2019 Nobel Prize in Physics. As both a fan of history and a practicing cosmologist, I found his book to be a captivating read that immediately put my understanding of the universe into a new historical perspective. It should be a mandatory read for cosmology graduate students and seasoned cosmologists alike.

Peebles begins his book by discussing Albert Einstein's key ideas: that a philosophically sensible universe was homogeneous and isotropic, and that in general relativity such a universe must expand or contract. He then reviews the discovery of the expansion of the universe, evidenced by the redshift of distant galaxies, and the establishment of Big Bang cosmology. Throughout the book, Peebles carefully highlights alternative research directions that appeared promising at first but eventually fell out of favor due to their conflicts with observations. He emphasizes the historical importance of those false starts, such as the once-popular steady-state cosmology, and argues that they were catalysts for the development of innovative ideas that tested what eventually became the dominant paradigm.

Peebles then outlines how the abundance of light elements and, most importantly, the discovery of the cosmic microwave background (CMB) tilted the balance in favor of the hot Big Bang model. It is fascinating to see how George Gamow had the amazing intuition to establish that picture decades before it came into focus observationally. A description of early ideas about the formation of structure and galaxies in our universe follows. The author's depiction of early hypotheses about the formation of galaxies is somewhat reassuring: It shows that scientists first tried the most intuitive approach—which failed—and only arrived at today's more complex models after decades of refinement.

Cosmology's Century contains an illuminating and riveting discussion of the growing evidence for subluminal matter in astronomy during the 20th century. It gradually became apparent that the mysterious subluminal matter was identical to the nonbaryonic dark matter proposed by particle physicists. The text also tracks the numerous theoretical and experimental dead ends that eventually led to the community's acceptance of the cold dark matter (CDM) paradigm. Peebles himself was instrumental in that process: In 1982 he wrote a paper showing that CDM provides a simple explanation for both the highly inhomogeneous state of the local universe and the smoothness of the CMB. Showing characteristic humility, he recounts his surprise at how quickly the model became a convincing picture of the early universe. Yet again, Peebles's insightful speculation led the community toward a fuller cosmological picture.

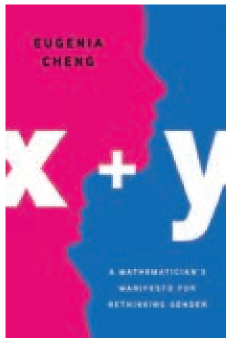
The book ends with an account of the cosmological "revolution" of 1998–2003, which firmly established the validity of the Lambda–Cold Dark Matter (ΛCDM) cosmological model. That such a simple model can explain both the initial seconds after the Big Bang and the distribution of galaxies today is a testament to a century of cosmology. Those of us who entered the field after that revolutionary era can find it easy to take our current cosmological paradigm for granted and forget that it took decades to get there.

Cosmology's Century serves as a reminder that we stand on the shoulders of giants. It also holds many lessons for today's researchers. At one point, Peebles emphasizes the importance of speculation to scientific progress, humorously saying that "one of the arts of science is to probe the boundaries between empty and productive speculation." He also draws a historical parallel to Lord Kelvin's turn-of-the-century pronouncement that "clouds" were hanging over physics. Those "clouds" led to the discovery of special relativity and quantum mechanics. Peebles similarly sees two "clouds" hanging over the ΛCDM model of the universe: the unknown physics of the very early universe and the enigmatic simplicity of dark energy and dark matter. Will those "clouds" lead to genuinely new ideas in physics?

Peebles accurately points out that large failures akin to the breakdown of classical physics on the atomic scale have yet to occur in cosmology. He hopes that such failures will lead to a much deeper understanding of our universe. Will the next century teach us as much about cosmology as the last?

Francis-Yan Cyr-Racine
University of New Mexico
Albuquerque

NEW BOOKS & MEDIA

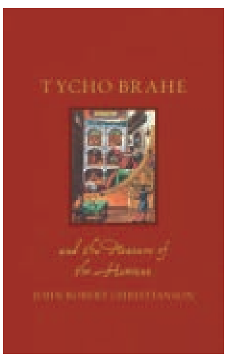


x + y

A Mathematician's Manifesto for Rethinking Gender

Eugenia Cheng
Basic Books, 2020. $28.00

Women face more barriers in science, technology, engineering, and mathematics than men. To get at the root of the problem, Eugenia Cheng explores personal characteristics in her new book, *x + y*. She argues that society values ingressive behaviors, such as being competitive and independent, more than congressive ones, such as being collaborative and caring toward others. But if we focus less on whether those characteristics have masculine and feminine connotations and instead redefine successful people to include those with predominantly congressive behaviors, the representation of women and other underrepresented groups in science may improve. –AL



Tycho Brahe and the Measure of the Heavens

John Robert Christianson
Reaktion Books, 2020. $22.50

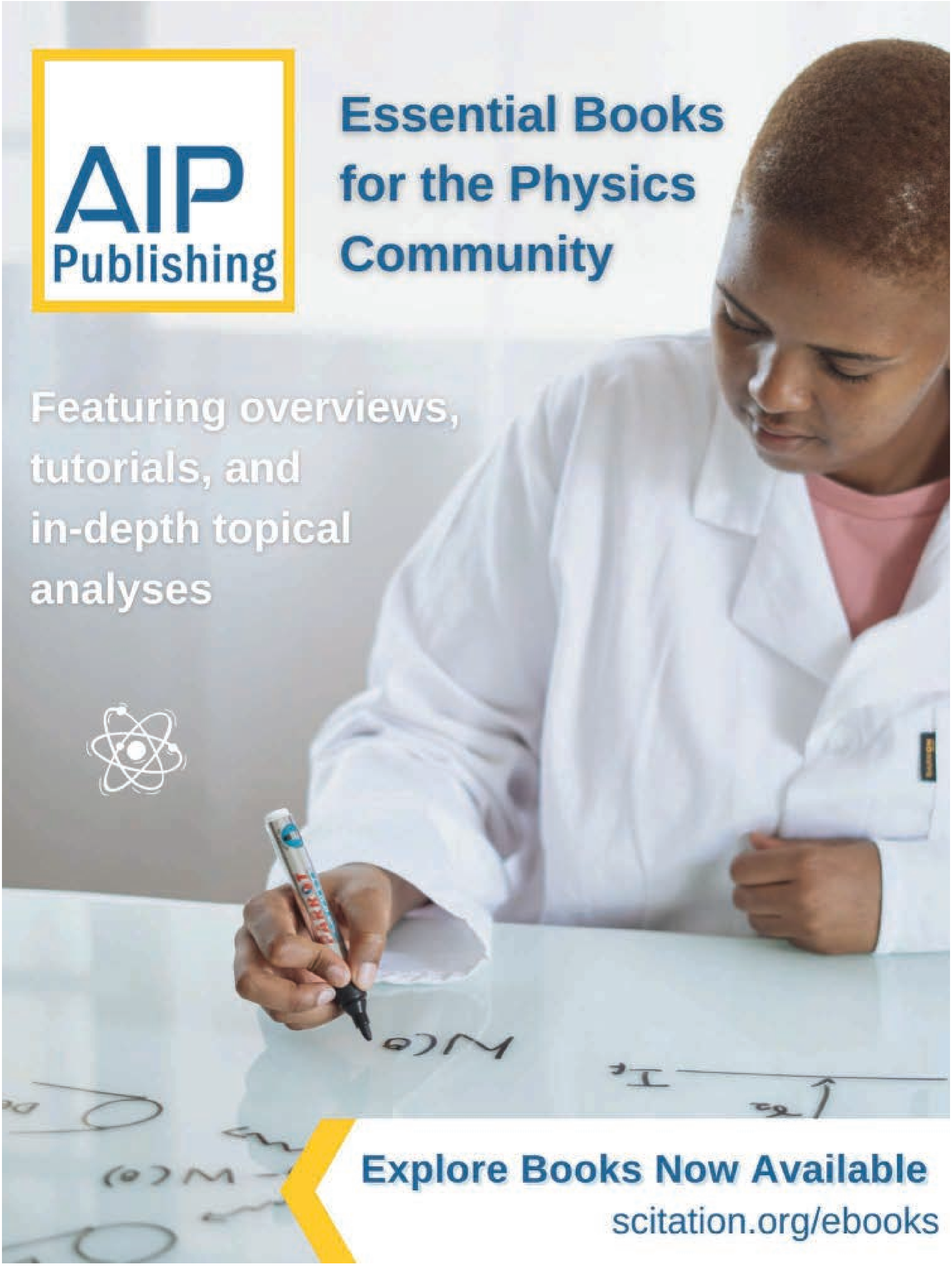
A new biography aims to restore early modern Danish astronomer Tycho Brahe to what the author, historian John Robert Christianson, asserts is his rightful place as a "world-class figure of the late Renaissance." Richly illustrated and based on a detailed analysis of Tycho's surviving texts, *Tycho Brahe and the Measure of the Heavens* also draws on significant new scholarship written in the Nordic languages. Christianson deftly depicts Tycho's tumultuous relationship with his one-time assistant Johannes Kepler, while paying careful attention to the context of Tycho's position in the Danish and Bohemian courts and the sociopolitical ramifications of the Reformation. Ultimately, Christianson argues that Tycho's empirical methods had a "profound and enduring" effect on the history of science. The book is required reading for scholars of early modern astronomy. –RD

Anti-vaxxers

How to Challenge a Misinformed Movement

Jonathan M. Berman
MIT Press, 2020. $19.95 (paper)

Written by renal physiologist and science educator Jonathan Berman, *Anti-vaxxers* is a how-to guide for those looking to counter rising anti-vaccine sentiment across the globe. Nevertheless, much of the book describes the intertwined history of vaccinations and anti-vaccine movements; surprisingly, Berman shows that anti-vaccine sentiment has existed since the first vaccine was developed in the 19th century. Today the movement is fueled by social media, "fake news," and other disinformation. Berman argues that "community-based" strategies that consider the identity and values of parents and groups targeted by anti-vaxxers are the most likely to succeed in convincing them of vaccine safety. The book will be of interest to science-minded parents and policymakers alike. –RD PT


AIP Publishing
Essential Books for the Physics Community
Featuring overviews, tutorials, and in-depth topical analyses
Explore Books Now Available
scitation.org/ebooks

NEW PRODUCTS

Focus on software, data acquisition, and instrumentation

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

Andreas Mandelis

Optical power meter

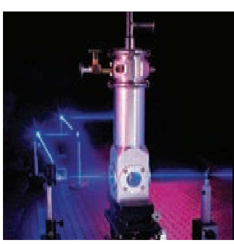
The compact, cost-effective F-712.IRPx photometer from Physik Instrumente (PI) is designed mainly for use in aligning glass fibers, fiber arrays, and other elements to photonically integrated circuits. The optical power meter detects optical input signals, without range switching, in the 600–1700 nm wavelength range with a power of 230 pW to 1.3 mW. The signals are converted with high resolution into a precise logarithmic voltage signal of 0.1–1.6 V. The signal strength changes depending on the geometrical position of the elements to be coupled, and the change detected by the power meter is used as a control variable for position optimization. The 6 kHz bandwidth enables the high-throughput alignment and scanning capabilities of PI's fast multichannel photonic alignment (FMPA) motion subsystems, solution platforms, and microrobots. Two models of the F-712.IRPx are available: one for integration into PI's FMPA systems and the other for use as a single-channel benchtop or panel-mounted device. ***Physik Instrumente LP****, 16 Albert St, Auburn, MA 01501, www.pi-usa.us*

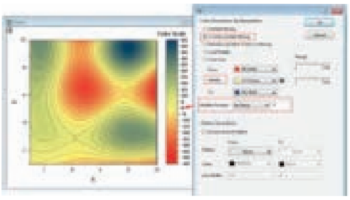
Electromagnetic compatibility test software

AR RF/Microwave Instrumentation has updated its electromagnetic compatibility (EMC) test software to make EMC testing more complete, reliable, and economical. For ease of use, version 5.0 of AR's emcware offers accelerated startup and one-click reporting. Its streamlined user workflow, more than 500 built-in testing profiles, prebuilt reporting templates, efficient software validation, and reverb immunity testing methodologies help reduce setup time and training. To meet the varied needs of users involved in EMC testing, AR has also made available Nexio BAT-EMC test software, which offers enhanced flexibility for EMC and non-EMC testing and for in-depth equipment management. ***AR RF/Microwave Instrumentation****, 160 Schoolhouse Rd, Souderton, PA 18964, www.arworld.us*

Mathematical programming software

MathWorks has introduced version R2020b of its MATLAB and Simulink software. New capabilities in MATLAB simplify working with graphics and apps. Simulink expands access and increases speed, and the launch of Simulink Online enables the viewing, editing, and simulating of Simulink models through Web browsers. R2020b also introduces new products that build on artificial intelligence capabilities, speed up autonomous systems development, and accelerate creation of 3D scenes for automated driving simulation. Among new and updated features in MATLAB are bubble and swarm charts, the ability to diff and merge App Designer apps with the MATLAB Comparison Tool, and customizable figure icons and components in MATLAB apps. R2020b can generate code for referenced model hierarchies in Simulink up to two times as fast as the previous version. It also offers new automerge functionality that helps automate continuous integration workflows. ***The MathWorks Inc****, 1 Apple Hill Dr, Natick, MA 01760-2098, www.mathworks.com*


JANIS
A LAKE SHORE COMPANY
sales@lakeshore.com
www.lakeshore.com/janis
Does your research require low temperatures? Contact Lake Shore today. Our engineers will assist you in choosing the best system for your applications.



Data analysis and graphing software

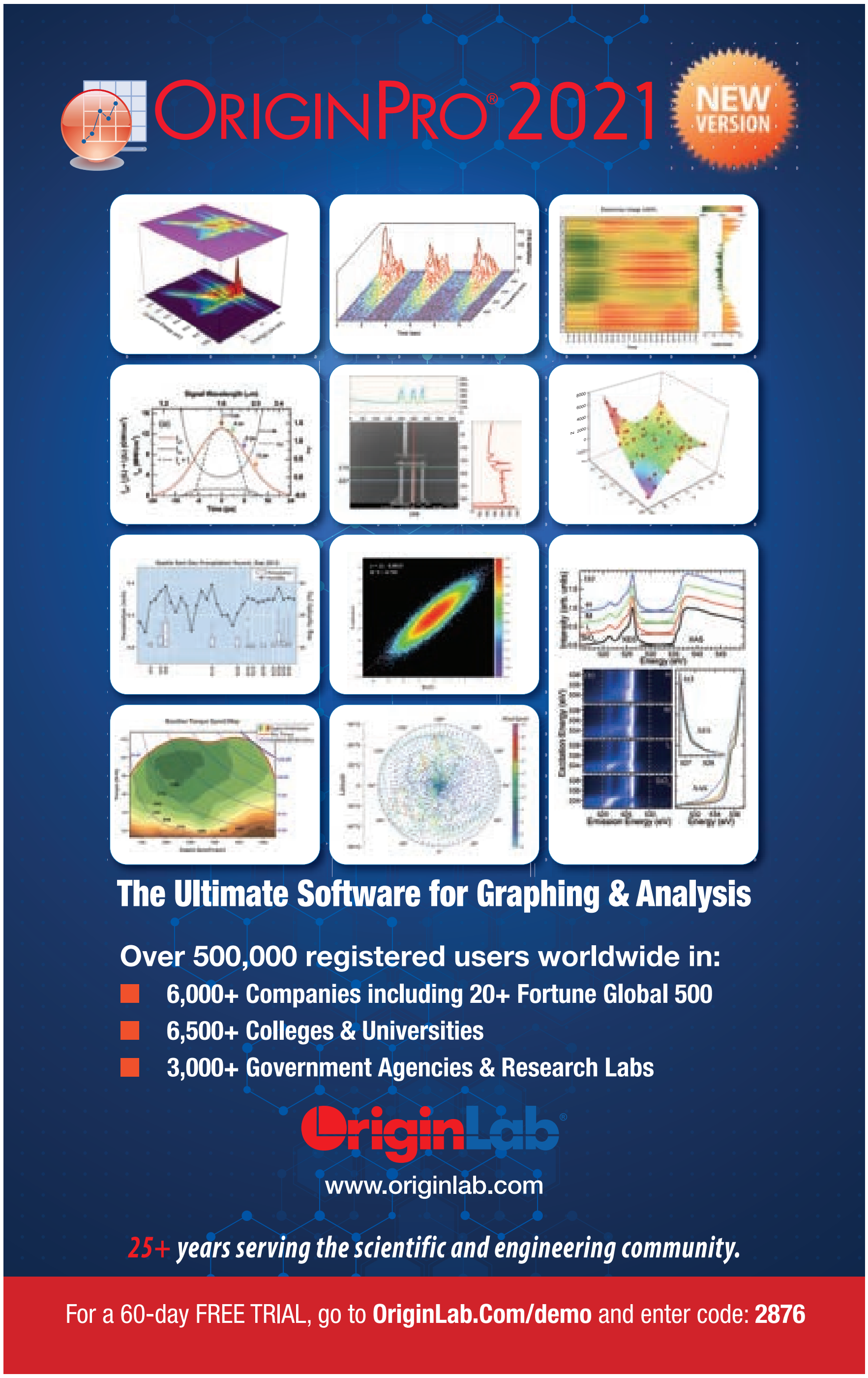
Version 2020b of Origin and OriginPro, OriginLab's data analysis and graphing software, adds more than 75 new features, apps, and improvements. Mini toolbars for worksheets and matrix sheets now give users quick access to common operations at the cell, column, row, and worksheet levels. Mini toolbars have more buttons for manipulating page, layer, and plot settings, and they can be customized by adding and removing buttons. Browser graph templates allow for quick, easy, multichannel data exploration. Enhancements include a new Data Navigator panel, the ability of users to connect to multiple files, and improvements to existing MATLAB, HDF, Excel, and NetCDF data connectors. New graph types include Bland–Altman and bee-swarm plots, and new apps include 3D antenna radiation pattern and x-ray diffraction analysis. ***OriginLab Corporation***, *One Roundhouse Plaza, Ste 303, Northampton, MA 01060, www.originlab.com*

Crystal structure software

CrystalMaker has released the latest iteration of its software for visualizing crystal and molecular structures in research and teaching in chemistry, solid-state physics, materials science, mineralogy, and crystallography. In addition to miscellaneous enhancements and bug fixes, version 10.5.4 includes a revised set of van der Waals radii. The revision, based on recently published data, provides more comprehensive coverage with more accurate values. With the addition of Bond Range Generation for Text Files, CrystalMaker text files can now include new "BRNG" cards, which specify automatic bond—and hence polyhedra—generation between a specified pair of element symbols, within the specified distance range between the pair. According to the company, such cards provide a more precise alternative to the existing "BOND" or "BMAX" cards. ***CrystalMaker Software Ltd***, *Centre for Innovation & Enterprise, Oxford University Begbroke Science Park, Woodstock Rd, Begbroke, Oxfordshire, OX5 1PF, UK, www.crystalmaker.com*

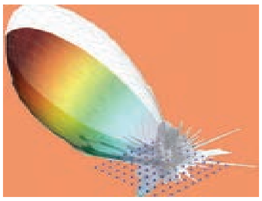
Three-axis magnetic field sensor

Bartington Instruments has announced its CryoMag fluxgate magnetometer, which provides high-precision measurements of static and alternating magnetic fields. The three-axis cryogenic-compatible sensor offers noise performance of less than 20 pT_{rms} at 1 Hz. It has three concurrent axes, a 100 μT measuring range, and a frequency response from DC to 1 kHz (±5%). It easily connects via USB to the new DG-1 digitizer. The CryoMag can operate at temperatures down to 2 K, which makes it suitable for use in cryostats for superconducting RF cavities and magnetic field monitoring, quantum computing applications, and other low-temperature environments and applications. It can also be employed as a feedback sensor in active magnetic field cancellation systems. ***Bartington Instruments Ltd***, *10 Thorney Leys Business Park, Witney, Oxfordshire OX28 4GG, UK, www.bartington.com*



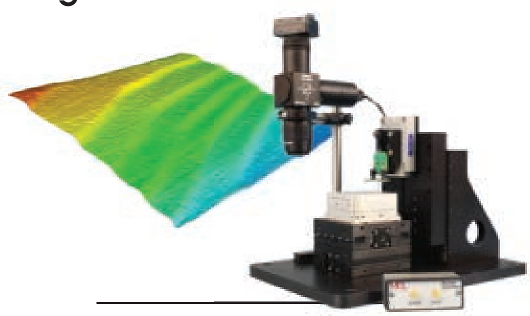
EM digital receiver and direction finder

Rohde & Schwarz has launched an economical, optimized receiver and direction finder with minimum size, weight, and power requirements. The R&S EM200 digital receiver has a wide frequency range and supports spectrum monitoring. It detects, analyzes, and demodulates from 8 kHz to 8 GHz. The EM200 is a suitable in-phase and quadrature source for continuous wideband data collection. Capable of remote monitoring and remote power cycling, it is an appropriate handoff receiver. It can be used in signal interception systems and in combination with R&S signal analysis software, recorders, and most receivers. The EM200 can be optionally equipped with the company's compact direction-finding antennas, which have a range from 20 MHz to 6 GHz. ***Rohde & Schwarz GmbH & Co KG***, *Muehldorfstrasse 15, 81671 Munich, Germany, www.rohde-schwarz.com*



Multiphysics simulation software

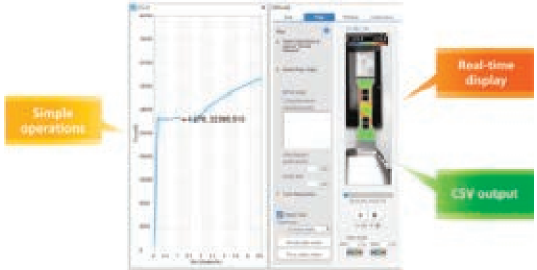
The latest version of the Tech-X multiplatform, multiphysics simulation software, VSim, is now available. It runs computationally intensive electromagnetic, electrostatic, and plasma simulations in the presence of complex dielectric, magnetic, and metallic shapes. VSim 10.1 has added not only Boolean operation functionality to computer-aided design (CAD) geometries but also the ability to duplicate CAD geometries into rectangular arrays. A new radiation-reaction feature in the VSimPA plasma acceleration package lets relativistic electrons emit photons that can be optionally tracked in the simulation. The simulation allows for recording of the emission energy (weight), position, and number of relativistic gamma photons. The VSimEM electromagnetics package now offers an example of phased array antenna simulation that lets users easily set up their own multiple-element antenna to analyze its radiation pattern and optimize the design. Other new example simulations in VSim 10.1 include the Antenna Array Single Excited Element and the Visual Setup Version of Electron-Beam Driven Plasma. ***Tech-X Corporation***, *5621 Arapahoe Ave, Ste A, Boulder, CO 80303, www.txcorp.com*

MCL®
MAD CITY LABS INC.
High Resolution AFM and NSOM

Atomic Step Resolution
Closed Loop Nanopositioners
Precalibrated Position Sensors
Integrated Z-axis Control
Automated Software Control
Designed for DIY AFM
sales@madcitylabs.com
www.madcitylabs.com

Large-sample AFM software

Oxford Instruments Asylum Research has reported that its Jupiter XR atomic force microscope (AFM) now includes the company's new Ergo software interface. According to Asylum Research, the Jupiter XR is the fastest and highest-resolution large-sample AFM available and the only one with blueDrive technology for high measurement stability and repeatability. Ergo's intuitive guided workflow makes it faster and simpler to set up measurements and collect high-quality AFM images in multiuser academic research and industrial R&D laboratories. Intelligent algorithms automatically calculate the optimal settings that will allow both expert and infrequent AFM users to achieve consistent, high-quality images. ***Oxford Instruments America Inc***, *300 Baker Ave, Ste 150, Concord, MA 01742, www.oxinst.com*

Real-time strain visualization software



Strain View software from Shimadzu Scientific Instruments (SSI) provides strain distribution visualization for precise materials testing measurements. It is easy to operate and integrates with the company's TRViewX digital video extensometer for simplified strain distribution imaging and data processing. Strain View software displays the strain distribution in real time, so users can observe it while conducting a test. Since video files are linked with the stress–strain curve, users can view strain distribution at specific points of interest. Because the software is fully integrated with SSI's Trapezium X-V software, users can view, capture, analyze, and report data from one platform. Strain data can be saved in CSV format and exported to simulation software. ***Shimadzu Scientific Instruments Inc***, *7102 Riverwood Dr, Columbia, MD 21046, www.shimadzu.com*

Compact gas analyzers

Pfeiffer Vacuum has unveiled compact, portable gas analyzers for quantitative and qualitative analysis of gases at atmospheric pressure in such areas as metallurgy and environmental analysis, chemical processes, and the semiconductor industry. The OmniStar and ThermoStar GSD 350 feature a low detection limit—depending on the mass range—of less than 100 ppb, low gas consumption of 1–2 sccm, and a fast measuring time of up to 1 ms/u. They allow simultaneous detection of all gases within the mass range. Their gas inlet, fitted with a heated capillary for use at up to 350 °C, prevents vapors from condensing during process gas analysis. The two-stage inlet system makes possible an almost segregation-free gas supply. The ThermoStar GSD 350 was developed for coupling with thermo balances. Its inlet system, which has a quartz capillary and a platinum orifice, enables analysis of even the smallest concentrations. Designed for a wide range of applications, the OmniStar GSD 350 has a stainless-steel capillary and a valve that can interrupt the sample gas stream. ***Pfeiffer Vacuum Inc***, *24 Trafalgar Sq, Nashua, NH 03063, www.pfeiffer-vacuum.com* PT

OBITUARIES

Harry Suhl

Harry Suhl

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Harry Suhl, who made seminal contributions to condensed-matter physics and was among the founding faculty at the University of California, San Diego, died on 3 March 2020 after an enviably long, happy, and productive life. He was also well known and much beloved as a quick wit, a gourmet, a sharp dresser, and an all-around bon vivant.

Harry was born in Leipzig, Germany, on 18 October 1922. In 1938, during a threatening period for German Jews such as Harry and his family, his father, Bernard, wrote to a tenuous British business contact requesting that they be sponsored for immigration as refugees. In May 1939 the Suhls emigrated from Leipzig to London, where Harry finished his secondary schooling.

A year later Harry and Bernard were taken to a series of internment camps, which the British government had established for aliens with potential Axis sympathies. At Huyton, near Liverpool, a community developed among the Jewish intelligentsia, who were permitted access to books and records and who sponsored public lectures on subjects such as the new "quantum theory." It was in that setting where Harry's interest in physics was first kindled. Following a string of happy accidents, he found himself released from internment and studying at the University of Wales. After earning his BSc in 1943, he worked on radar for the British Navy until he began his PhD studies in ionospheric physics at Oxford University in 1946.

By late 1948 Harry had his PhD and a job at Bell Labs. It was a remarkable period at the R&D company, especially for solid-state physics. The technical staff at the Murray Hill, New Jersey, campus was brimming with current and future luminaries, such as John Bardeen, Charles Kittel, Bernd Matthias, Philip Anderson, George Feher, and Theodore Geballe. Also, far less of a distinction was made between experimentalists and theorists at that time than now.

Suhl's earliest work at Bell Labs was on charge-carrier dynamics in semiconductors in magnetic fields. In 1953 Suhl and Larry Walker comprehensively analyzed wave propagation in waveguides that are filled with gyromagnetic and gyroelectric media, relevant to various microwave devices. In 1955–56 Suhl provided the definitive explanation of nonlinear effects in ferromagnetic resonance, now known as the Suhl instability. That work led to his getting a patent for a ferromagnetic parametric amplifier in 1956 and inspired wide use of parametric amplification in general. In 1957 Suhl and Tuto Nakamura independently uncovered a major source of broadening of NMR lines in magnetically ordered media, now known as the Suhl–Nakamura interaction.

Suhl's interests then turned to superconductivity, where he extended the Bardeen-Cooper-Schrieffer theory to two-band systems, and to quantum many-body problems. In 1961, shortly after arriving at the just-opened University of California, La Jolla—later renamed UC San Diego—he made major contributions to the theory of many-body effects on impurity states in metals. Suhl showed how the recently discovered Kondo singularity was replaced by a Fermi surface resonance, a feature now known as the Abrikosov–Suhl resonance, in dilute magnetic alloys. Throughout the remainder of his career, Suhl continued to work on various aspects of magnetism, but he also branched out into such areas as surface physics, catalysis and reaction kinetics, nonlinear dynamics, and biological physics.

Suhl served on the editorial board for *Physical Review* in 1963–72 and for *Solid State Communications* in 1961–90; was coeditor with George Rado of the highly influential five-volume series *Magnetism: A Treatise on Modern Theory and Materials* (Academic Press, 1963–73); and authored the 2007 monograph *Relaxation Processes in Micromagnetics* (Oxford University Press). He also twice served as chair of the UC San Diego physics department, and he was director of its Institute for Pure and Applied Physical Sciences from 1980 to 1991.

Among his friends and colleagues, Harry was regarded with deep affection for his wit and conviviality. When once asked what he did to keep fit, Harry replied, "Oh, I really don't subscribe to strenuous exercise. However, I do get up every morning and wind my watch by an open window." At a wonderfully inspiring and generous speech at his own 70th birthday conference, Harry offered the following "unwelcome advice to the younger people": "Above all, don't get wiser as you get older. If you do, you will become too inhibited to try the impossible, and one can achieve the limits of the possible only by occasionally venturing beyond them. The famous proverb should really be transposed: Angels rush in where fools fear to tread."

Daniel Arovas
M. Brian Maple
University of California, San Diego
La Jolla
Pradeep Kumar
University of Florida
Gainesville PT

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

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

RECENTLY POSTED NOTICES AT www.physicstoday.org/obituaries

Guido Münch
9 June 1921 – 29 April 2020

Otto Sankey
11 January 1951 – 21 March 2020

Alexander Patashinski
8 August 1936 – 22 February 2020

Jorge Ramiro Antillón Matta
13 April 1931 – 6 February 2020

Connecting physicists and engineers with the finest jobs

Find your future at
physicstoday.org/jobs

PHYSICS TODAY

QUICK STUDY

Marc Miskin is an assistant professor of electrical and systems engineering at the University of Pennsylvania in Philadelphia.

Making robots microscopic

Marc Miskin

Wirelessly powered and mass manufactured, the dust-sized robots can crawl, survive harsh environments, and be injected through a hypodermic needle.

The first time I looked into a microscope, I was amazed. Earlier that day, my science teacher had waded into the creek near our elementary school and pulled up a bucket of muddy water. After he distributed samples on microscope slides to the class, I brought the image into focus and found myself staring at a tiny monster—a rotifer, sucking up helpless microbes around it—trapped in a drop of water.

That memory is not unique. Most people are probably stunned the first (or really any) time they peer through a microscope. In spite of their small size, the creatures that populate the tiny worlds that come into view are remarkably complex. They can sense, communicate, process information, and explore. What if it were possible to build a machine the size of a microorganism that engineers could communicate with and directly control? I learned early on that nature builds such organic robots all the time. The question is how to pull it off synthetically for ourselves.

A new technology

Fortunately, scientists already know how to build many of the key components needed for a microscopic robot. Electronics have been shrinking for the past 50 years. And with today's state-of-the-art fabrication technology, nearly a million transistors can fit in the space of a paramecium. The smallest computer—complete with a microprocessor, memory, a thermometer for sensing, and an LED for two-way communication—is just a few times as large as a hair's width. In principle, those technologies could readily be adapted into the core of tiny machines that sense, think, and act.

That's the vision: Steal the powerful design of microelectronics and exploit it for robotics. For five years I worked toward that goal as a postdoc with Paul McEuen and Itai Cohen at Cornell University, and now I head my own group at the University of Pennsylvania. When we started, the biggest roadblock was clear enough: We could build the brain of a microrobot but not the body. Plenty of semiconductor electronic components were available, but no microscale actuators that could seamlessly integrate with them. The available approaches to building moving parts for microrobots encountered one of two problems: Either the parts couldn't be fabricated to bend without breaking at cellular dimensions or the actuators couldn't be controlled with electronic signals.

To overcome those hurdles, we designed a new technology: surface electrochemical actuators (SEAs). Each SEA is made from a thin layer of platinum—ours is 7 nm thick—capped on one side with a 2 nm layer of titanium. The entire actuator is just a dozen atoms thick, which makes it flexible enough to deform without breaking.

Put a SEA in water, apply a voltage to it, and you'll find that it bends. That's because aqueous ions attach themselves to the surface of the platinum to an extent that depends on the voltage. And when they attach, they stress the material, but only on the platinum side. The difference in stress between platinum and its capping layer prompts the device to curl. At the microscale, those surface forces are significant thanks to the large surface-to-volume ratio of tiny objects. What's more, the curling is reversible: When the voltage is turned off, the ions detach from the surface and the stress relaxes.

The voltage required to attach or remove ions is a few hundred millivolts. It's constrained by thermodynamics to be essentially the thermal energy per ion, kT/e, where k is the Boltzmann constant, T is temperature, and e is the charge on an electron. The same scale also determines the turn-on voltages for silicon microelectronics. As a result, integration is trivial: One can simply run a wire from the output of a silicon transistor into a SEA. And the same electrical signal used for computation can be used for mechanical work.

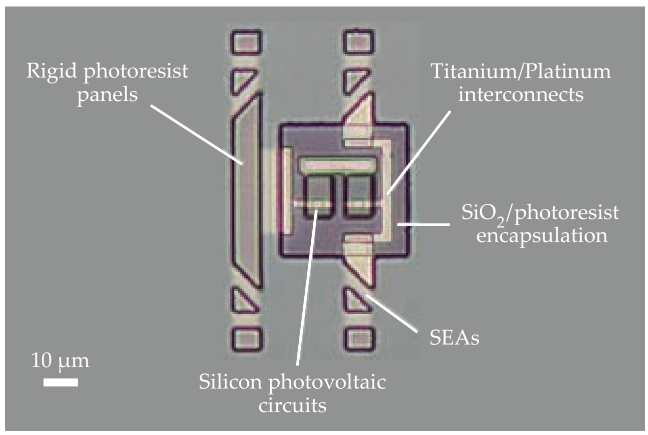


FIGURE 1. THIS MICROSCOPIC ROBOT is composed of a body with internal electronics and four movable legs. The circuits are made from silicon p–n junctions and metal interconnects, encapsulated between a layer of silicon dioxide and a layer of photoresist. The robot is also equipped with two tiny solar cells, the photovoltaic circuits. By shining laser light on each cell, researchers create the voltage that controls the front and back legs, which are made from a new class of surface electrochemical actuators, or SEAs, that allow the robot to move. (Adapted from M. Z. Miskin et al., *Nature* **584**, 557, 2020.)

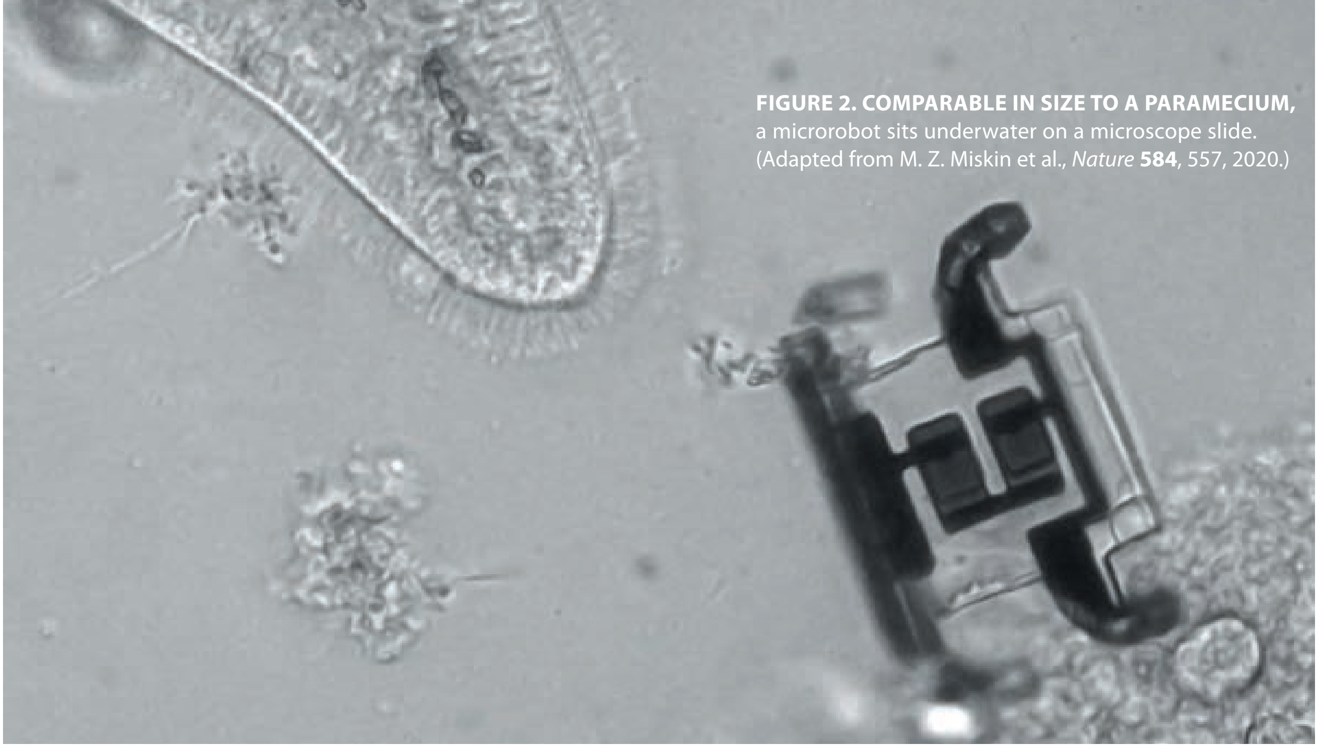


FIGURE 2. COMPARABLE IN SIZE TO A PARAMECIUM, a microrobot sits underwater on a microscope slide. (Adapted from M. Z. Miskin et al., *Nature* **584**, 557, 2020.)

From principle to proof

We built a first prototype about two years ago by linking electronics with actuators. The robot, shown in figure 1, is simple. Atop its body, it supports two silicon photovoltaic circuits that convert laser light into electrical energy. The photovoltaics power two sets of legs, one in the front of the body and the other in the back. Laser pulses shining on the photovoltaics force each set of legs to move on command. By repeatedly switching between moving the front and rear legs, we can make the robot walk around its environment. As far as we're aware, the device is the smallest-legged walking robot in existence. At just under 100 μm in length, it is too small to be seen by the naked eye (see figure 2).

Actually, we didn't build *a* prototype; we built 1 million. The technology used for microfabrication is massively parallel. We can literally make the million robots atop a 4-inch silicon wafer. We also figured out how to release them in parallel: We can reliably set the robots free into an aqueous environment to produce tiny robot swarms that float freely and are ready to walk. The largest swarm we've built so far had about 10 000 members in it.

Built millions at a time, microrobots wind up being extremely inexpensive, with each costing fractions of a penny. They're essentially disposable, a far cry from their expensive macroscopic cousins. Indeed, being tiny endows the robots with a host of unique properties. Walking at the microscale consumes just nanowatts. With solar cells and SEAs as actuators, the robots can be powered by laser light or even ordinary sunlight and walk around indefinitely. They can respond to the light signal in 10 ms—a limit imposed by the viscous drag of the fluid—and exert forces on the order of nanonewtons. That's on par with the forces biological cells generate during locomotion and more than enough to lift and move the 50 pN weight of the robot's body. Currently robots walk at about 1 μm/s, a speed set by how fast we switch the laser. Yet the peak speeds can be as fast as 30 μm/s—almost one body length per second.

The tiny machines are also remarkably robust. Their exterior is made of platinum, glass, and a durable polymer epoxy. The result is extreme chemical and thermal stability. We've exposed our robots to strong acids, strong bases, and temperature variations of several hundred degrees, and we've shot them through syringes. In each case, they survived.

An active field

Beyond our group, research teams around the world are working on other ways to build microscale machines. Some groups use magnetic fields, ultrasound, or chemical reactions to power and control robots. Some are working to adapt those advances to the medical field, so the robots can operate deep in the human body at single-cell resolution. The microworld is emerging as an environment that humans can now control, manipulate, and engineer.

A bright future awaits. Scientists can now adapt the new actuators into systems designed for locomotion, for grasping, and for manipulating the local environment. My own group is in the early stages of designing and building walking robots whose on-board circuitry responds to stimuli, times maneuvers, and is even capable of basic programmability.

We're starting to envision applications such as rewiring damaged nerves, cleaning battery electrodes, and using microrobots themselves as model systems for investigating the physics of biological swarms. If you're wondering what else is possible from building small machines, think back to your own experiences looking into a microscope. The microworld is the fundamental scale of biology. Even the simplest living things on Earth hint at the awesome opportunities for microscopic machines.

Additional resources

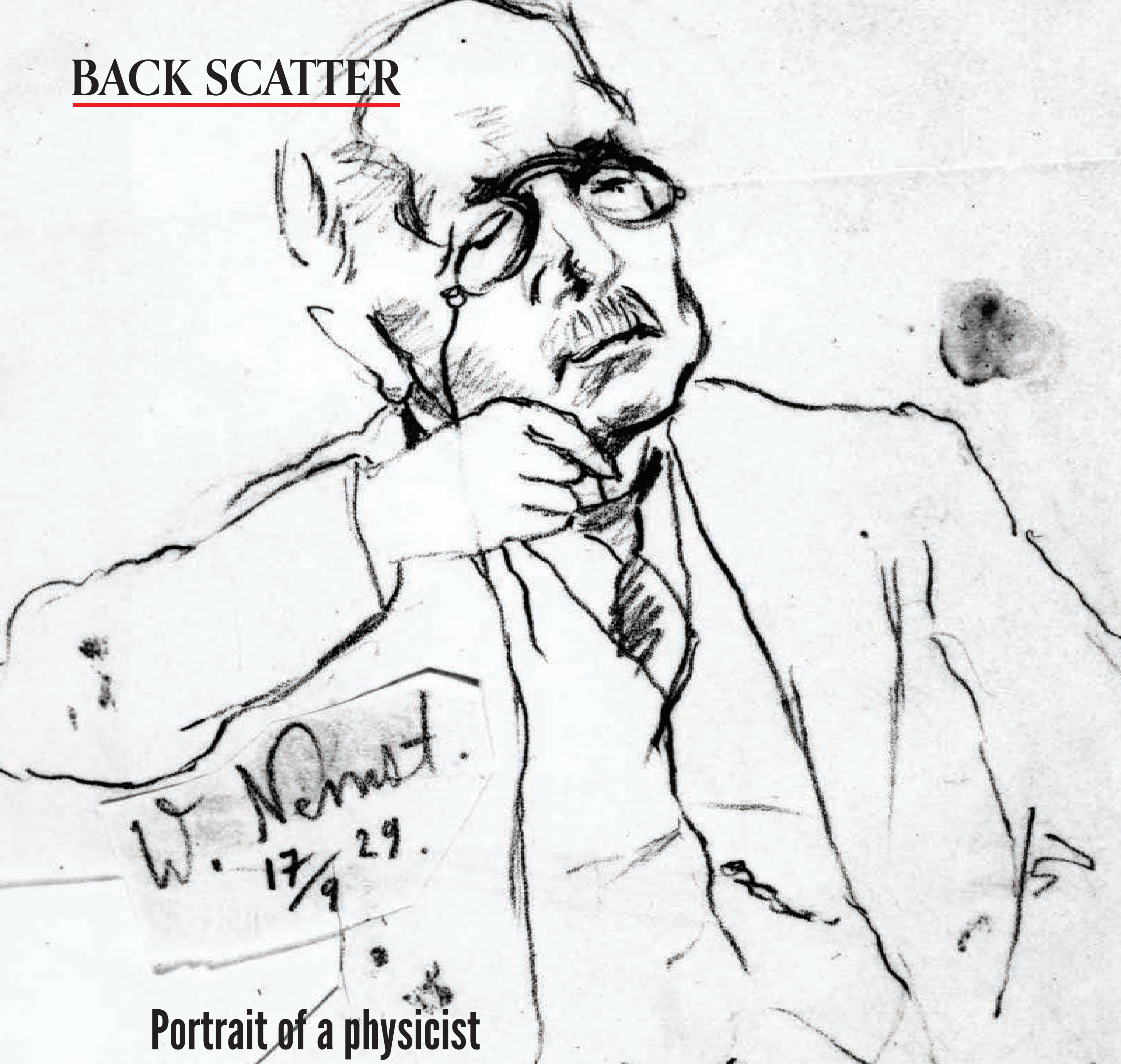
▶ M. Miskin, P. McEuen, "Tiny robots with giant potential," TED talk (November 2019).

▶ M. Z. Miskin et al., "Electronically integrated, mass-manufactured, microscopic robots," *Nature* **584**, 557 (2020).

▶ M. Z. Miskin et al., "Graphene-based bimorphs for micron-sized autonomous origami machines," *Proc. Natl. Acad. Sci. USA* **115**, 466 (2018).

▶ K. J. Dorsey et al., "Atomic layer deposition for membranes, metamaterials, and mechanisms," *Adv. Mater.* **31**, 1901944 (2019).

▶ M. F. Reynolds et al., "Capillary origami with atomically thin membranes," *Nano Lett.* **19**, 6221 (2019). PT



Portrait of a physicist

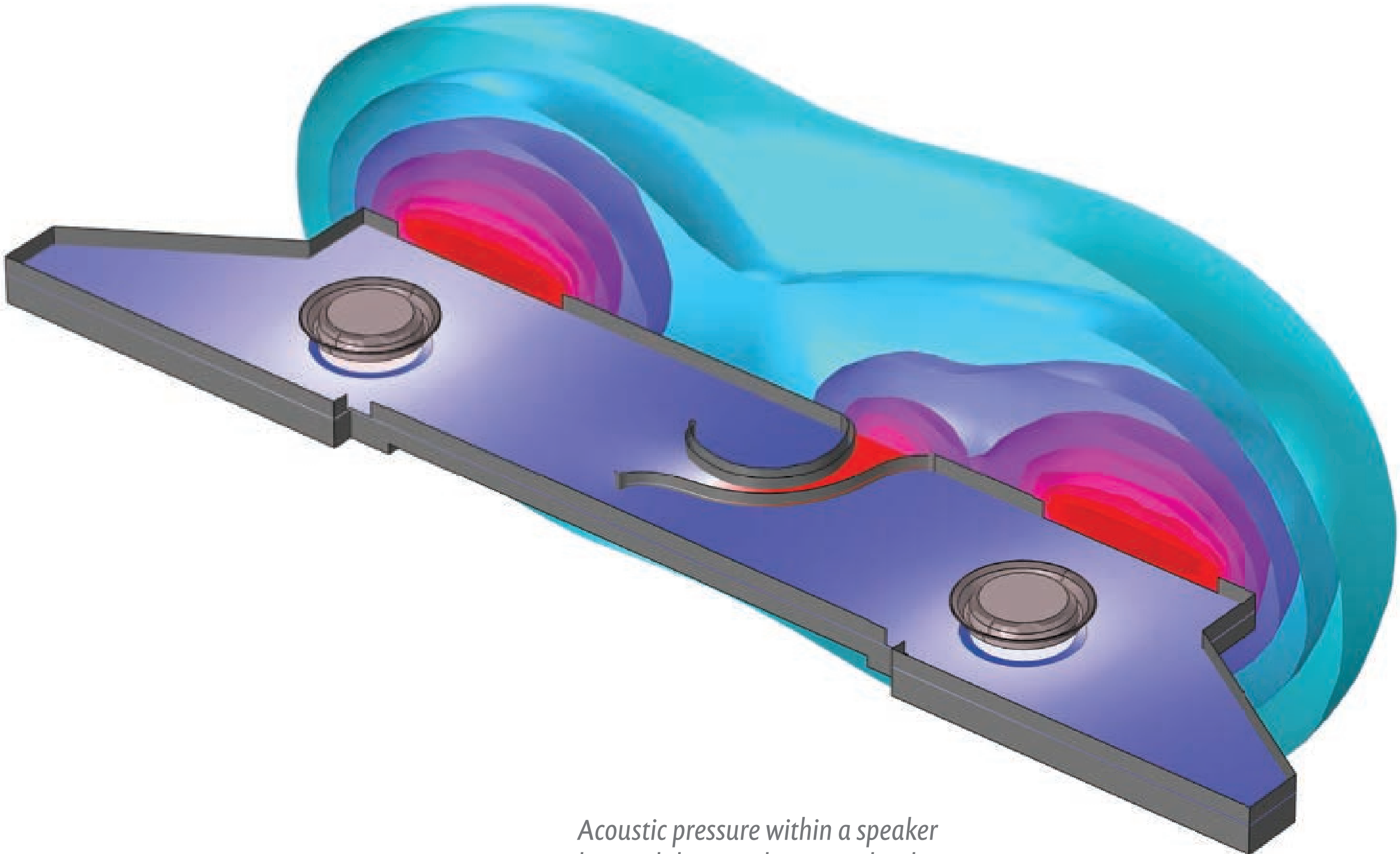
Elizabeth Korn, a Jewish German artist and children's book illustrator, drew portraits of several well-known European scientists in the 1920s and 1930s. She sketched her subjects during conferences and meetings that she attended with her husband, Arthur, a physicist and inventor. Pictured here is one of her sketches of the German chemist Walther Nernst. He devised the Nernst heat theorem, which states that as a system approaches absolute zero, so, too, does the entropy change of any chemical reaction or physical transformation.

In 1939 Korn and her husband fled Nazi Germany and eventually settled in New Jersey. She was a professor of art at Drew University, where she worked until her retirement in 1966, and she became known for her abstract expressionist mixed-media artwork. Korn's collection of portraits were acquired by the Niels Bohr Library and Archives at the American Institute of Physics (publisher of PHYSICS TODAY) in 1999. The Emilio Segrè Visual Archives there currently holds about 100 portraits of several notable scientists she sketched, including Niels Bohr, Marie Curie, Otto Hahn, Werner Heisenberg, and Max von Laue. Additional collections of Korn's work are at the Smithsonian Archives of American Art and the Museum of Modern Art.

—KJ

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

Simulation + testing = optimized loudspeaker designs



Acoustic pressure within a speaker box and the sound pressure level in the surrounding domain.

A global leader in electronics rose to the top of the audio industry by adding multiphysics simulation to their design workflow. COMSOL Multiphysics® enables audio engineers to couple acoustics analyses and other physical phenomena to address design challenges inherent to loudspeaker and soundbar designs.

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

comsol.blog/loudspeaker-design

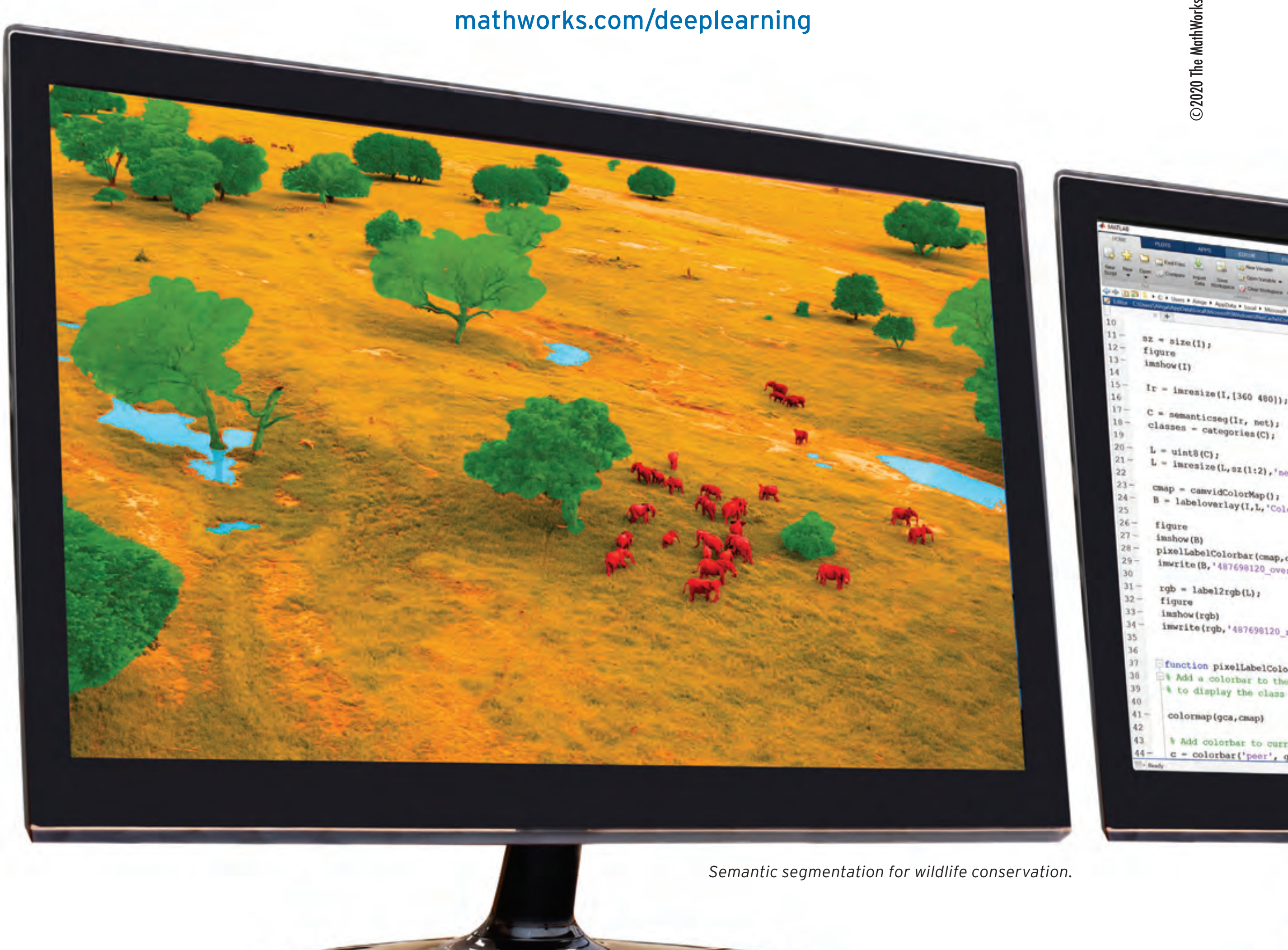
COMSOL

MathWorks®

MATLAB SPEAKS DEEP LEARNING

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

mathworks.com/deeplearning



Semantic segmentation for wildlife conservation.

©2020 The MathWorks, Inc.