PHYSICS TODAY

December 2021• volume 74, number 12

A publication of the American Institute of Physics

Soft matter

Branched flows

Quantum startups in Israel

Starving star formation

We're with you all the way

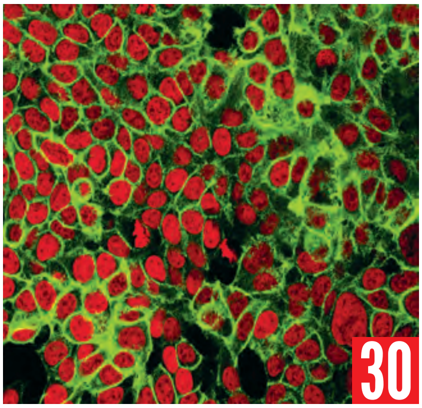
The **broadest bandwidth** with the **highest power.**

AR offers over **100 amplifiers** ranging from **10 Hz - 50 GHz** and with power levels of **1 W - 100 kW.**

PHYSICS TODAY

December 2021 | volume 74 number 12

FEATURES



ON THE COVER: The blue-ringed octopus's distinguishing characteristics owe their bright color to the self-assembly of colloidal particles. On **page 38**, Bas van Ravensteijn and Ilja Voets describe colloidal materials whose self-assembly out of equilibrium requires light or molecular fuel to sustain it. On **page 30**, Amanda Parker describes how the soft tissues of humans and other organisms can be regarded as mechanical metamaterials. Together, the two articles provide a taste of the vitality and variety of the physics of soft matter. (Photo by kaschibo/Shutterstock.com.)

Recently on PHYSICS TODAY ONLINE

www.physicstoday.org

MARTIN WHITE

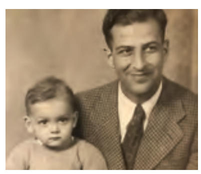
Joanne Cohn

Before there was arXiv, there was Joanne Cohn. As a postdoc in the late 1980s, she started an informal exchange of string theory manuscripts that eventually became the physics preprint server. Three decades later, researchers submit some 16 000 manuscripts to arXiv each month and download 30 million.
physicstoday.org/Dec2021a

BLACK IN PHYSICS

#BlackInPhysics Week

Physics Today and *Physics World* recently copublished the second annual #BlackInPhysics Week essay series. Five Black physicists detailed their experiences with burnout, a critical topic as Black scholars confront systemic racism both within and outside of academia.
physicstoday.org/Dec2021b



TITO PONTECORVO

Tito Pontecorvo

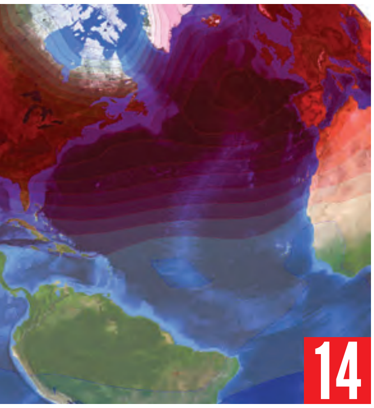
In 1950 nuclear physicist Bruno Pontecorvo defected with his family to the Soviet Union. His son Tito, six years old at the time, was later forced out of his job as an oceanographer by the Communist Party. In a Q&A, Tito tells his family story and explains how in 1997 he left the country for the US with 75 horses.
physicstoday.org/Dec2021c

Physics Today (ISSN 0031-9228, coden PHTOAD) volume 74, 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.

PHYSICS TODAY

www.physicstoday.org

DEPARTMENTS



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

QUANTUM MATERIALS, QUANTUM INFORMATION, AND QUANTUM COMPUTING

The Departments of Physics and Astronomy, Materials Science & Engineering, and Electrical Engineering and Computer Science at the University of Tennessee, Knoxville, invite applications for three tenure-track assistant professor level positions. These positions are part of an interdisciplinary cluster hire in the area of Quantum Materials for Future Technologies, and are expected to begin August 1, 2022. Successful applicants may have joint appointments between the relevant departments.

The positions are:

1. A computer scientist or physicist with interest and expertise in one or more of the following areas: quantum information, quantum computation, quantum materials, artificial intelligence, machine learning. Preference will be given to candidates who can interface between these areas; **https://apply.interfolio.com/95968**.

2. A condensed matter experimentalist with interest and expertise in scanning tunneling microscopy and spectroscopy. Preference will be given to researchers who can interface with quantum materials researchers and other members of the clusters; **https://apply.interfolio.com/95969**.

3. A researcher with interest and expertise in experimental quantum materials research with a focus on device fabrication and nanoscale transport phenomena. Preference will be given to candidates who can interface with the other members of the cluster; **https://apply.interfolio.com/95952**.

The successful candidate will join a thriving interdisciplinary research community working in the broader area of quantum materials and quantum information, including three recently appointed cluster faculty. The cluster aims to accelerate the development of quantum science, artificial intelligence, and their technological applications, and to develop interdisciplinary courses in this area.

Candidates should have a Ph.D. in physics, materials science, electrical engineering, computer science or related field. The successful candidate has an exemplary record of research with significant contributions to their field of study as demonstrated through peer-reviewed publications, invited talks, leadership roles, and the candidate's Statement of Research. The successful candidate will be expected to establish an externally funded research program and provide training for graduate students and postdocs. He or she is furthermore committed to effective undergraduate and graduate teaching, curricular innovation, and creative teaching methods, as demonstrated by their Teaching Statement.

The University respects and values people of all races, genders, creeds, cultures, and sexual orientations, and holds a deep commitment toward developing and promoting an inclusive community. We strongly encourage applications from members of groups that are underrepresented in STEM fields, as well as candidates who will contribute in meaningful ways to the equity and inclusion goals of the three departments, as demonstrated in the candidate's Statement on Diversity, Equity, and Inclusion.

UT Knoxville, is Tennessee's flagship university, a campus of choice for outstanding undergraduates, and a premier graduate institution. Its researchers benefit from the available resources and proximity to unique research facilities at Oak Ridge National Laboratory.

The campus is in one of the most beautiful areas of the country, with easy access to miles of inland waterways, pristine state and national parks, diverse cultural opportunities, and a blend of convenient urban and rural living settings. Downtown Knoxville, adjacent to campus, is a thriving neighborhood filled with restaurants, shops, and indoor and outdoor entertainment venues, and is home to the Knoxville Symphony Orchestra, the Knoxville Opera Company, annual international festivals, such as the Big Ears Music Festival, and seasonal weekly events such as Jazz on the Square, Shakespeare on the Square, and the Farmers Market.

Applicants should send a 1) *Curriculum Vitae* with a list of publications and invited talks, 2) Statement of Research with proposed research program, 3) Statement of Teaching, 4) Statement of Diversity, Equity, and Inclusion, and arrange for at least three letters of reference to be submitted separately. All application materials should be submitted via the Interfolio links provided above. Acceptable file formats are .pdf or .docx. Applications reviews will start January 10, 2022.

FROM THE EDITOR

It's all too much

Charles Day

This past October, Johan Chu of Northwestern University and James Evans of the University of Chicago published their analysis of bibliometric data drawn from 10 broad fields in science, medicine, engineering, and mathematics.[1] Their main conclusion is sobering: As the number of papers in a field increases, researchers find it harder to recognize innovative work. Progress seems to be slowing.

Chu and Evans present a plausible hypothesis. Rather than evaluate new papers on their individual merits, researchers increasingly resort to comparing them with existing paradigms. What's more, when new papers are published at a high rate, truly novel ideas struggle to prevail—let alone be noticed—over the flood of competitors.

A strength of Chu and Evans's paper is that they used their hypothesis to make six predictions, all of which they could test by looking for correlations. The predictions are, to quote the paper,

> 1) new citations will be more likely to cite the most-cited papers rather than less-cited papers; 2) the list of most-cited papers will change little year to year—the canon ossifies; 3) the probability a new paper eventually becomes canon will drop; 4) new papers that do rise into the ranks of those most cited will not do so through gradual, cumulative processes of diffusion; 5) the proportion of newly published papers developing existing scientific ideas will increase and the proportion disrupting existing ideas will decrease; and 6) the probability of a new paper becoming highly disruptive will decline.

Citations are straightforward to count. To characterize a paper's canonicity, disruptiveness, and diffusibility, Chu and Evans developed statistical measures. In all, they examined 1.8 billion citations of 90 million papers published from 1960 to 2014. Each of their six predictions was affirmed by a significant correlation.

Physics was represented in Chu and Evans's study by applied physics. Could averaging progress across such a broad field yield spurious correlations? I wondered. At the start of their range, 1960, the laser had just been invented, and transistors had yet to supplant vacuum tubes in electronic devices. At the middle of their range, 1987, high-T_c superconductivity had recently been discovered, and magnetic tape remained a popular storage medium. Research activity in those and other areas of applied physics waxed and waned with fashion, but I couldn't see how those fluctuations might mimic a secular rise in, say, the canon's rigidity. It's also hard to see how correlations that began to be noticeable in the 1970s could be attributed to China's impressive and more recent growth in scientific output.

One alternative explanation for the correlations could be that the most-cited papers are not the revolutionary paradigm shifters that Chu and Evans presume. In 2014 Richard Van Noorden, Brendan Maher, and Regina Nuzzo published their analysis of the 100 most cited papers in science.[2] Top of their list was "Protein measurement with the Folin phenol reagent," which in 1951 introduced what became known as the Lowry protein assay.[3] Like other top papers, it describes a useful, widely applicable method. When the number of scientists increases, so does the number of citations to papers that describe such methods.

But that explanation doesn't hold up, at least not in physics. The most cited papers in *Applied Physics Letters* are not method papers. Indeed, one of them helped Shuji Nakamura win a share of the 2014 Nobel Prize in Physics for developing the blue LED.[4]

If Chu and Evans's conclusions are valid—and I think they are—then what is to be done? Telling physicists to publish less, publishers to stop launching new journals, and editors to reject more papers are all illiberal restrictions on freedom of expression. But scientists need a better way to evaluate papers' novelty. Whoever develops one would earn scientists' gratitude. They might even become rich.

REFERENCES

1. J. S. G. Chu, J. A. Evans, *Proc. Natl. Acad. Sci. USA* **118**, e2021636118 (2021).
2. R. Van Noorden, B. Maher, R. Nuzzo, *Nature* **514**, 550 (2014).
3. O. H. Lowry et al., *J. Biol. Chem.* **193**, 265 (1951).
4. S. Nakamura, T. Mukai, M. Senoh, *Appl. Phys. Lett.* **64**, 1687 (1994).


INTRODUCING...
GradSchoolShopper
MAGAZINE
The Student Guide to Grad School in Physics, Astronomy, and Related Fields
READ NOW!
gradschoolshopper.com/magazine
Choosing the right program for you
Make your application stand out
Applying to grad school with a COVID-19 transcript
presented by
AIP | American Institute of Physics

The harsh truth about terraforming

I enjoyed reading Charles Day's column on the subject of terraforming in the July 2021 issue of PHYSICS TODAY (page 8). I was only a year old in 1942, so I missed seeing Jack Williamson's story that introduced the term when it was published. I also missed it in the mid 1950s when I read many of the back issues of *Astounding Science-Fiction*.

My first introduction to the concept was from reading *Sands of Mars* by Arthur C. Clarke, first published in 1951, although I don't believe he actually used the term "terraforming." He describes selectively breeding native plants that extract oxygen from the Martian soil to release it into the atmosphere—a precursor of the biological approach in Kim Stanley Robinson's Mars trilogy—and turning Phobos into a second sun to warm the planet.

On a more serious note, it seems to me that the very concept of terraforming represents a striking display of human hubris. Today we are struggling to safeguard the health of Earth's critical ecosystems from the unintended consequences of the artificial environment we have built. Given our limited success so far and significant chance of failure, the idea that we could create a viable, self-sustaining environment on another planet anytime soon seems pretty far-fetched. The reality is that a planetary ecosystem is many orders of magnitude more complex than what we are currently able to deal with. The time will likely come when our capabilities will have reached that level, but for the moment, terraforming appears destined to remain in the realm of science fiction and, apparently, board games.

MARS'S SURFACE as captured by NASA's *Curiosity* rover. The image's white balance has been adjusted to show how the surface of Mars would look under Earth's skylight. Terraforming Mars has been the subject of several works of science fiction, such as *Sands of Mars* by Arthur C. Clarke. (Courtesy of NASA/JPL-Caltech/MSSS.)

Robert Dickinson
(bob@argosanalytics.com)
Menlo Park, California

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.

More on Arrhenius plots

Axel Lorke's Quick Study in the May 2021 issue of PHYSICS TODAY (page 66) describes Svante Arrhenius's illustrious career and provides important insight into Arrhenius's quantitative description of thermally induced processes. Lorke describes the broad power of the famous Arrhenius relationship $a = A\exp(-E_a/k_BT)$, where k_B is the Boltzmann constant and T is absolute temperature, to capture complex physics with temperature-dependent measurements aimed at the extraction of a single parameter, the activation energy E_a. Presenting an example from thermodynamics and another from kinetics, he shows how E_a connects closely to independently determined quantities such as a semiconductor's bandgap energy and the UV-induced gelation energy of proteins.

The Quick Study focuses on the slope of logarithmic plots of rates and other temperature-dependent quantities versus inverse temperature $1/T$. In kinetics, the prefactor A of the exponential also provides important physical information. It may be obtained by extrapolating an Arrhenius line like that of figure 3 in Lorke's Quick Study to yield an intercept at the

$1/k_BT = 0$ axis. If the quantity measured is the frequency of a process, as is often the case in solid-state physics, the prefactor A can be called the attempt frequency, with A^{-1} being the limiting time required to surmount the activation barrier as the temperature approaches infinity. In textbook examples, for small E_a, this approach yields plausible values for such frequencies. Further, if the E_a of such a process is modified only slightly, the intercept does not change.

Starting with reports by Frederick Hurn Constable[1] in 1925 and by Wilfried Meyer and Hans Neldel[2] in 1937, researchers have done a great number of experiments on sets of closely related materials and systems in which the prefactor of Arrhenius plots of a related set varies systematically with E_a. While care must be taken to avoid artifacts, it has been clear for some time that the phenomenon is real.[3] For a wide-ranging variety of sets of related physical, geological, biological, and chemical phenomena, the logarithms of those intercepts vary linearly with E_a. That also means the Arrhenius fit lines cross at an isokinetic temperature at which the rate is independent of E_a. Those observations have various names: the isokinetic rule, the compensation law (because the increase in the prefactor partially compensates for the increase in E_a), and the Meyer–Neldel rule.

The meaning and explanation of the Meyer–Neldel rule were long considered to be a mystery, but work by a number of groups in the final decades of the past millennium provided a clear theoretical framework for both kinetic and equilibrium systems. The key to activation is not the energy or enthalpy; it is the free-energy change, which includes an entropy term. When the activation barrier is large, the entropy change increases with E_a, and that increases A.

In 2006 one of us (Yelon) coauthored a review of the state of the art in experiment and theory,[4] which have continued to evolve since. Systematic studies yield information concerning the characteristic energy of the collective excitations—phonons or local vibrations—that are aggregated to surmount the activation barrier.[5] In some cases, notably studies of electronic or ionic conductivity, important information concerning mechanisms can be obtained. Like the Arrhenius relation that spawned it, the Meyer–Neldel rule is an elegant way to gain insight into the fundamental interactions governing temperature-dependent processes.

References

1. F. H. Constable, *Proc. R. Soc. London A* **108**, 355 (1925).
2. W. Meyer, H. Neldel, *Phys. Z.* **38**, 1014 (1937).
3. W. Linert, *Collect. Czech. Chem. Commun.* **55**, 21 (1990).
4. A. Yelon, B. Movaghar, R. S. Crandall, *Rep. Prog. Phys.* **69**, 1145 (2006).
5. A. Yelon, B. Movaghar, H. M. Branz, *Phys. Rev. B* **46**, 12244 (1992).

Howard Branz
(howard@branztechnologypartners.com)
Branz Technology Partners
Boulder, Colorado

Arthur Yelon
(arthur.yelon@polymtl.ca)
Polytechnique Montréal
Montreal, Quebec, Canada

Axel Lorke's Quick Study (PHYSICS TODAY, May 2021, page 66) discusses underlying connections between seemingly unrelated phenomena. His brief reminder of Svante Arrhenius's major contributions to science—and in particular, his now-famous empirical relation describing the rate of thermally activated processes—is thought-provoking.

There is, however, a minor point appearing in the caption of figure 2, showing electron density versus inverse temperature, that needs clarification. If we concentrate on the high-temperature (left) portion of the graph, the "reaction" involved is the generation of an electron–hole pair (analogous to electron–positron pair production in particle physics), for which the Gibbs free-energy change is what is known as the energy gap E_g. For that reaction, the product of the electron density n and hole density p is given by $np \propto \exp(-E_g/k_BT)$, where k_B is the Boltzmann constant and T is absolute temperature.[1] At the high temperatures on the left side of the graph, the semiconductor is nearly intrinsic, so $n \approx p \propto \exp(-E_g/2k_BT)$, which is where the factor of two in the denominator of the slope originates.

The caption states that the "factor of two in the denominator arises because electrons obey Fermi–Dirac statistics rather than classical Boltzmann distribution." But at such high temperatures, Fermi–Dirac statistics approximates the Boltzmann distribution, so it cannot be the reason for the factor of two as the caption states. A similar explanation applies to the ionization of neutral donors.[2]

References

1. C. D. Thurmond, *J. Electrochem. Soc.* **122**, 1133 (1975).
2. R. W. Christy, *Am. J. Phys.* **40**, 40 (1972).

Remo A. Masut
(remo.masut@polymtl.ca)
Polytechnique Montréal
Montreal, Quebec, Canada

Another way to define physics

In his letter "But what is physics?" (PHYSICS TODAY, April 2021, page 12), Peter Zimmerman raises an interesting question about physics and physicists. He recalls the definition given by his former professor Leonard Schiff: "Physics is whatever physicists do."

Schiff's definition, however, seems like a tautology. I prefer one given by Gabriel Weinreich:

> Physics is delineated, not by its subject matter, but by the methods of thought that a physicist uses.[1]

Additionally, in their December 1995 article in PHYSICS TODAY (page 25), Sol Gruner, James Langer, Phil Nelson, and Viola Vogel provide a definition of physicist that I still treasure:

> The physicist is most cogently identified, not by the subject studied, but by the way in which a subject is studied and by the nature of the information being sought.

Both definitions point to the methods involved.

Reference

1. G. Weinreich, *Am. J. Phys.* **61**, 1067 (1993).

Julian Ting
(juhilian@gmail.com)
De-Font Institute for Advanced Studies
Taichung, Taiwan PT

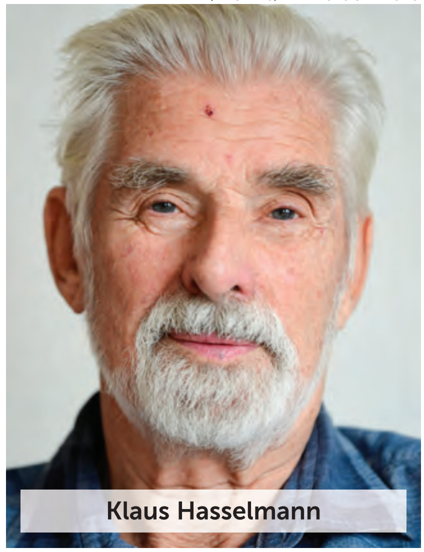
Climate modeling innovators are honored with half the physics Nobel

Syukuro Manabe pioneered the simulation of Earth's climate. Klaus Hasselmann extracted climate trends from noisy weather data and identified patterns of human-induced warming.

BENGT NYMAN/CC BY 2.0

Syukuro Manabe

FABIAN BIMMER/REUTERS/ALAMY STOCK PHOTO



Klaus Hasselmann

In 1946 John von Neumann at the Institute for Advanced Study in Princeton, New Jersey, began working on a project with meteorologist Carl-Gustaf Rossby: to predict the weather using one of the first programmable computers. Kristine Harper, a former meteorologist and now a historian of science at the University of Copenhagen in Denmark, says, "What von Neumann really had in mind was being able to control the weather, and the military was all over that."

Jule Charney joined the project in 1948; he and von Neumann made some simplifying assumptions about the atmosphere's density and disregarded vertical air motion to make the simulation more tractable. The result was the first numerical weather forecast, produced in 1950. Although it predicted a few things correctly, it was also beset with problems. Nevertheless, the rudimentary forecast inspired Joseph Smagorinsky, a US Weather Bureau meteorologist, to adapt the new numerical weather model for studying the climate.

To help with the task, Smagorinsky hired Syukuro Manabe, a recently minted PhD who arrived in Washington, DC, from Japan in 1958. Smagorinsky led the development of a three-dimensional model, and Manabe and his new colleagues got to work constructing a model of the atmosphere that accurately simulated radiative and convective processes. To assess the sensitivity of the atmosphere's temperature to carbon dioxide, they simulated a doubling of its concentration and found that the surface temperature increased by 2.3 °C, which is close to today's best estimate of 3 °C documented in the sixth assessment report of the Intergovernmental Panel on Climate Change (IPCC).[1]

In the 1970s Klaus Hasselmann was working at the Max Planck Institute for Meteorology in Hamburg, Germany, to better analyze climate model results. He developed a stochastic model to simulate the semi-random, chaotic weather in climate models. Later he devised a statistical technique for finding the anthropogenic warming trend in climate data. The variability driven by human action, such as the midlatitude warming in figure 1, has a pattern distinct from that of natural climate variability, which arises from fluctuations in the coupled ocean–atmosphere system.

For their foundational research in modeling and analyzing the complex system that is Earth's climate, Manabe and Hasselmann were awarded half of this year's Nobel Prize in Physics from the Royal Swedish Academy of Sciences. The other half was given to Giorgio Parisi for his work in understanding disordered systems (see page 17 of this issue).

"I'm very happy that they put the attention on the climate problem," said Hasselmann in an interview with Nobel Prize Outreach's chief scientific officer shortly after the announcement. Manabe told PHYSICS TODAY, "I never expected to receive this prize. Looking at past winners of the Nobel Physics Prize, they did outstanding research in fundamental physics. Now they [the Nobel committee] may start awarding prizes in our field of Earth science, and I am very encouraged."

Radiative-transfer model

The greenhouse effect of CO_2 on Earth's atmosphere was first recognized in the mid 19th century (see "Eunice Newton Foote's nearly forgotten discovery," by Maura Shapiro, PHYSICS TODAY online, 23 August 2021). Despite that early discovery, how Earth's climate might be changing as a result of the CO_2 that people were adding to the atmosphere failed to attract much serious attention. Many physical scientists argued that the ocean absorbed nearly all the additional CO_2. Other scientists felt that the actions of humans were too insignificant to sustain a geological effect as consequential as changing Earth's climate.

That mindset began to change in the mid 20th century. In 1953 Gilbert Plass took advantage of the advances in IR spectroscopy made during World War II to predict that more IR radiation would be absorbed in the atmosphere as more CO_2 was added. Using oceanographic measurements, scientists determined that it takes on average 10 years for CO_2 in the atmosphere to be absorbed by the ocean. (For more about the early history of climate science, see the article by Spencer Weart, PHYSICS TODAY, January 1997, page 34.)

Aware of Plass's work, Manabe set himself the goal of simulating the greenhouse effect and Earth's climate. He started with the planet's radiation budget. Under radiative equilibrium conditions, incoming shortwave solar radiation is balanced by Earth's outgoing longwave IR radiation. Manabe's model used a few inputs—incident solar light, Earth's average surface reflectivity, and the absorptivity of a few greenhouse gases, among others—and solved for the 1D vertical temperature profile of the atmosphere.

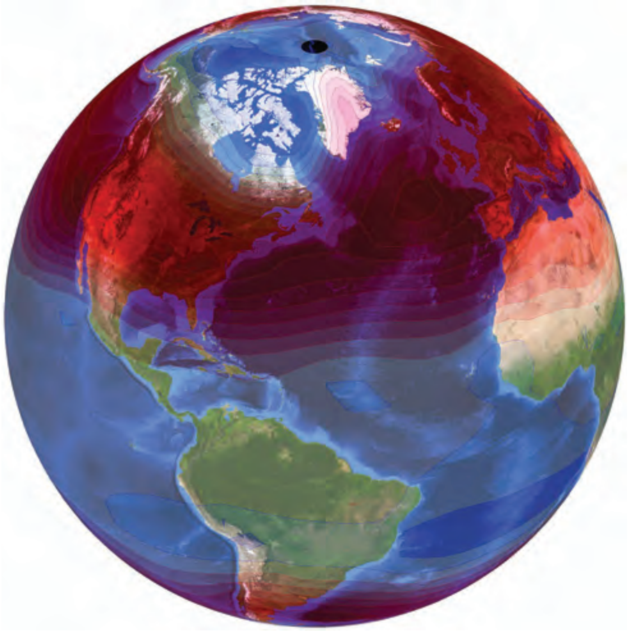


FIGURE 1. ANTHROPOGENIC WARMING is characterized by, among other spatial patterns, a significant increase in midlatitude air temperature relative to the tropics. More intense solar radiation near the equator generates temperature and pressure gradients that transport excess heat poleward, as shown here by the concentrated red banding overlying northern Africa, Europe, and North America. Klaus Hasselmann was the first to recognize that unique spatial patterns could be used to detect and attribute the causes of climate change. (From B. D. Santer et al., *Science* **361**, 245, 2018.)

In a 1961 paper, Manabe and his collaborator Fritz Möller, a German meteorologist who visited the US from 1959 to 1960, calculated the vertical temperature profile of the atmosphere for various regions and seasons.[2] Despite the simple, 1D representation of the atmosphere, the first results roughly agreed with observations.

Convective adjustments

The vertical temperature profile in Manabe and Möller's radiative equilibrium model lacked any adjustment that arises from the convective motion of the atmosphere. In a column of air with no convection, the temperature profile varies only by radiative transfer mediated by the greenhouse effect, at a rate of −15 °C/km.

In the real world, however, water vapor absorbs heat near Earth's surface and condenses as it rises. That phase change helps drive vertical motion in the atmosphere and effectively transfers thermal energy from the surface in the form of latent heat. The observed rate is about −6 °C/km.

To account for the effect of water vapor in the radiative equilibrium model, Manabe and Robert Strickler, a meteorologist at NOAA's Geophysical Fluid Dynamics Laboratory (GFDL), developed a convective adjustment that set the simulated lapse rate to the observed lapse rate.[3] "There were some pure radiative equilibrium calculations that people had done before, which don't really get you in the ballpark because they're not the right basic physics of the atmosphere," says Nadir Jeevanjee, a research scientist at the GFDL. "Manabe imposed the temperature profile corresponding to convection, which we now know to be the right assumption."

In 1967 Manabe and Richard Wetherald put all the pieces together to create a first-of-its-kind radiative–convective equilibrium model.[4] They allowed the absolute humidity—that is, the total mass of water in a given volume of air—to fluctuate over time. "When you warm the atmosphere, you want to let the absolute amount of moisture increase because that's what it tends to do in the real world," says Jeevanjee. "That was an innovation of the paper."

The critical result of the paper was its calculation of the climate sensitivity—an estimate of how much the atmosphere warms if the concentration of CO_2 doubles, shown in figure 2. The value of 2.3 °C is a bit lower than today's model average of 3 °C. The discrepancy may be, at least in part, from the Manabe and Wetherald model's lack of snow, sea ice, and other positive feedbacks in the climate system.

Ronald Stouffer's career at the GFDL overlapped with that of Manabe's for about 20 years. He says that "one of [Manabe's] many talents was his intuition, or scientific judgment. He was able to determine what you could throw out and what you should keep when you were developing a model."

Analyzing climate models

While Manabe and other scientists were developing more realistic climate models, fundamental advances were being made in the understanding of weather. Edward Lorenz, a meteorologist who spent most of his career at MIT, found in 1963 that the results of a simulation, whether it's modeling the atmosphere or some other physical system, can diverge widely for slightly different initial conditions and lead to unpredictable end states (see the article by Adilson Motter and David Campbell, PHYSICS TODAY, May 2013, page 27).

Hasselmann knew of Lorenz's work and wanted to better incorporate the chaotic nature of weather into climate simulations. He wrote an influential paper in 1976 introducing a model that simulates the climate probabilistically.[5]

The stochastic model takes air temperature, wind velocity, and a few other climate variables and uses them in a state vector to describe a climate system. A set of prognostic equations then uses the

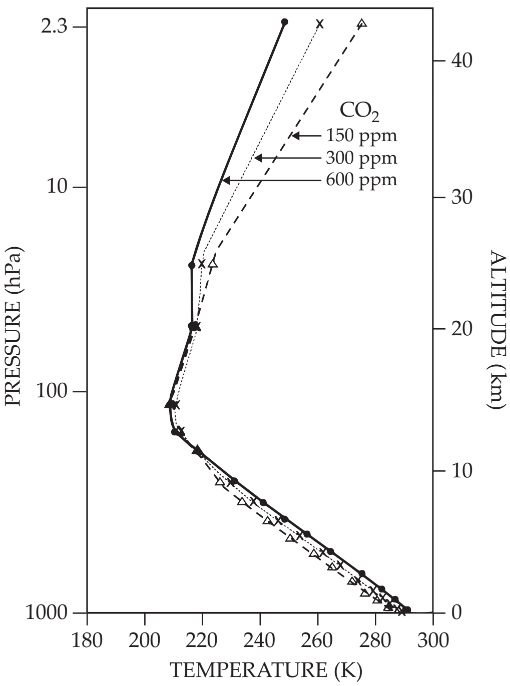


FIGURE 2. AIR TEMPERATURE varies with pressure, or altitude, in the atmosphere and according to the concentration of various greenhouse gases. Syukuro Manabe and Richard Wetherald simulated atmospheric temperature in a simple radiative–convective equilibrium model and in 1967 reported the first quantitative results for how much the atmosphere warms when carbon dioxide doubles. (Adapted from ref. 4.)

values of those variables at some initial time to predict their value at future times. Hasselmann—trained in physics and mathematics—also added to his model a form of the Fokker–Planck equation, a partial differential equation that describes random forces and drag acting on particles. He used the equation to represent chaotic weather perturbations.

Hasselmann then turned his attention to looking in simulations for specific signals, particularly the human-induced global warming trend. Previously, scientists had compared the average surface temperature at a location with the time-varying temperature over the course of the simulation to assess whether there was any statistically significant change.

The location-based approach was challenging because of the magnitude and time scale of climate change. The signal of interest—that is, the gradual monotonic increase in Earth's surface temperature—is almost always below the threshold of weather or natural variability in the climate system. On any given day, for example, weather conditions can cause the air temperature at some location to fluctuate by tens of degrees.

"Hasselmann's insight was to look at the entire pattern," says Benjamin Santer, an atmospheric scientist who recently retired from Lawrence Livermore National Laboratory. One such spatial pattern, or fingerprint, is that human-induced warming should manifest as a warmer lower atmosphere, where greenhouse gases are emitted, relative to a cooler upper atmosphere, where heat is lost to space. Observations and simulations, including Manabe and Wetherald's 1967 work,[4] have yielded that fingerprint many times over.

A 1979 paper by Hasselmann, which Santer calls "the beginning point of the entire detection and attribution field" (see PHYSICS TODAY, June 2021, page 19), stressed the importance of optimal detection.[6] That means looking not necessarily at where or when a signal of interest is strongest but where the noise is lowest or where and when the signal-to-noise ratio is highest.

Models show that the largest surface temperature change is occurring in the Arctic because of ice feedbacks. But the large natural variability there means that a more optimal fingerprint is the surface temperature change in tropical regions, where the natural variability is relatively small.

Hasselmann, Santer, Gabi Hegerl, and others in the field used the approach to identify changes in climate as measured in air temperature, ocean heat content, and other geophysical variables. When asked about working with Hasselmann, Hegerl recalls that "he was a great mentor, who encouraged us to be innovative, critical, curious, and think carefully about science, and he really enjoyed doing science, a joy that was contagious."

Curiosity and fun

Since its founding in 1988, the IPCC, a body of the United Nations, has asked whether and how humans are causing climate change. The models first built by Manabe and the analysis techniques based on Hasselmann's work have made it possible to answer that question with ever more accuracy and precision (see the article by Spencer Weart, PHYSICS TODAY, September 2015, page 46).

In 1990 the first IPCC report stated that more information was needed, but by 1995 the success of fingerprinting analyses led the authors of the second report to conclude that the balance of evidence suggested a discernible human influence on climate. In its most recent report, released a few months ago, the IPCC concludes by consensus that the human fingerprint is "unequivocal."[1] "To me," says Santer, "that is all traceable back to Hasselmann, the 1979 paper, and his admonition to young scientists like me to look at patterns."

Anthony Broccoli, a Rutgers University professor and long-time collaborator of Manabe, says of him that "what's most remarkable is that at the age of 90, the enthusiasm for science and his excitement in talking to people about science is undiminished." In January 2020 Manabe and Broccoli published a book titled *Beyond Global Warming: How Numerical Models Revealed the Secrets of Climate Change*, which details how Manabe came to understand the climate system (see PHYSICS TODAY, September 2020, page 54).

On the day of the Nobel Prize in Physics announcement, Manabe said at a Princeton University press conference, "I never imagined that this thing I would begin to study would have such huge consequences. I was doing it just because of my curiosity."

Alex Lopatka

References

1. V. Masson-Delmotte et al., eds., *Climate Change 2021: The Physical Science Basis—Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge U. Press (in press).
2. S. Manabe, F. Möller, *Mon. Weather Rev.* **89**, 503 (1961).
3. S. Manabe, R. F. Strickler, *J. Atmos. Sci.* **21**, 361 (1964).
4. S. Manabe, R. T. Wetherald, *J. Atmos. Sci.* **24**, 241 (1967).
5. K. Hasselmann, *Tellus* **28**, 473 (1976).
6. K. Hasselmann, in *Meteorology Over the Tropical Oceans*, D. B. Shaw, ed., Royal Meteorological Society (1979), p. 251. Available at https://pure.mpg.de/rest/items/item_3030122_2/component/file_3030123/content.

Parisi shares Nobel Prize for breakthroughs in understanding disordered systems

The laureate's solution to a spin-glass model underlies techniques used in a variety fields, including neuroscience and computer engineering.

The development of mathematical tools for understanding many-particle systems was one of the greatest achievements of 19th-century science. Extracting macroscopic properties of a gas by tracking the random motions of each individual atom remains an impossible feat. Statistical mechanics provided a framework for properly averaging over the individual particles' motions and understanding how they lead to collective macroscopic properties such as temperature and pressure.

But disorder takes many forms. Whereas gas molecules bouncing around a bottle sample the phase space of possible positions and velocities, the molecules that make up that glass or plastic bottle are frozen in disarray, trapped far from equilibrium. Explaining how those disordered molecules ended up stuck in a particular configuration and how that connects to any macroscopic properties is currently beyond the reach of statistical mechanics.

Giorgio Parisi is to receive half of this year's Nobel Prize in Physics for finding order in seemingly disordered complex systems. Among other significant contributions, he solved a model of a particular type of disordered system, spin glasses.[1] "By attacking the problem analytically and being able to solve it, he found a different paradigm for what ordered phases can be," says Leticia Cugliandolo, a professor at Sorbonne University in Paris. The underlying structures he found have helped researchers tackle seemingly intractable problems in neural networks, hard-sphere jamming, and protein folding, among other research areas.

A frustrating puzzle

Dilute magnetic alloys were created in the 1950s to experimentally study interactions between magnetic ions. They consisted of a few magnetic atoms dispersed in a nonmagnetic material—for example, manganese in zinc oxide. When impurity levels were low, the materials behaved as expected. But when the concentration of magnetic ions reached a few percent, measurements of the materials' magnetic susceptibilities and specific heats as functions of temperature yielded features that suggested a phase transition.

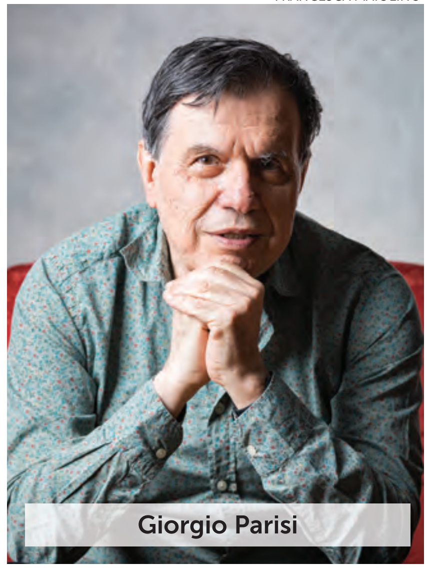
As in a ferromagnet, a magnetic alloy's spins become frozen in a particular configuration below a certain critical temperature T_c. But whereas a ferromagnet's spins are aligned, those in alloys were seemingly random (see the article by Daniel S. Fisher, Geoffrey M. Grinstein, and Anil Khurana, PHYSICS TODAY, December 1988, page 56). That frozen disorder led to the alloys being called spin glasses.

The randomness in spin orientations arises because the ions' positions are randomly distributed and their interactions can be ferromagnetic or antiferromagnetic. Those constraints make it impossible to minimize all the interaction energies simultaneously, and the spins are said to be frustrated (see the article by Roderich Moessner and Art Ramirez, PHYSICS TODAY, February 2006, page 24). Without a best configuration, the system ends up in one of the many seemingly random good-enough options.

Accounting for the properties of spin glasses piqued the interest of theorists, notably Philip Anderson and Sam Edwards. (Anderson wrote a seven-part column in PHYSICS TODAY about the history of spin-glass research that includes more detail than this story; the first installment appeared in the January 1988 issue on page 9.) Compared with other disordered systems, such as structural glasses and polymers, spin-glass systems seemed as though they might be amenable to a simple and solvable model because their disorder is baked into their fixed interactions.

In 1975 Anderson and Edwards used the random Heisenberg Hamiltonian, $H = \sum J_{ij}\, \mathbf{S}_i \cdot \mathbf{S}_j$, to describe an arrangement of spins $\mathbf{S}_i$ that can point in any direction.[2] The interaction strengths J_{ij} decrease with separation and, importantly, can take on

FRANCESCA MAIOLINO



Giorgio Parisi

both positive and negative values, thereby encompassing both ferromagnetic and antiferromagnetic spin pairings.

Calculating the system's thermodynamic properties required averaging the logarithm of its partition function Z over many sample configurations. To do it, Edwards devised the replica method, which involved taking the average of many versions of the system—replicas—that have the same interaction matrix **J**. The model yielded a cusp in the magnetic susceptibility at the phase transition, in agreement with experiments. It also produced features not seen in experiments, including a cusp in the specific heat.

To gain more insight, David Sherrington and Scott Kirkpatrick simplified the model further.[3] They gave the spin interactions infinite range, thereby creating a mean-field model, and replaced the Heisenberg spins with Ising ones that could point in only two directions, up and down.

The simplification hoisted a glaring red flag: negative entropy. Says Sherrington, "There is a known pathology for classical continuous variables that when you go to zero temperature, the entropy can become negative." The negative entropy seen at zero temperature with Heisenberg spins should have disappeared with discrete Ising spins. The persistence of negative entropy in the

Sherrington–Kirkpatrick model indicated a serious problem.

Unexpected structure

The culprit was replica-symmetry breaking.[4] Edwards and Anderson had assumed, correctly, that below T_c a particular spin's orientations in different replicas α and β would be correlated—meaning $q_{\alpha\beta} = \langle \mathbf{S}_i^\alpha \cdot \mathbf{S}_i^\beta \rangle$ was nonzero for $\alpha \neq \beta$. But they also presumed that any two replicas would yield the same result, $q_{\alpha\beta} = q$, after they averaged over all the spins. In fact, the situation was more complicated.

How the symmetry was broken, however, remained unclear. "The problem was hanging around with various people trying to think, What had we done wrong?" says Sherrington, a professor emeritus at Oxford University. "And that's where Parisi came out of the blue with this fantastic new idea, which solved the problem." His mathematical approach, published[1] in 1979, fixed the negative entropy issue. It also enabled predictions of spin-glass properties, such as magnetic susceptibility, that turned out to be correct.

"He found a mathematical solution of the problem, much more advanced than the previous one," says Hélène Bouchiat, a director of research at the Laboratory of Solid-State Physics, a joint research unit of the CNRS and the University of Paris-Saclay. "Mathematicians, I think, thought he was completely crazy." To calculate the system's free energy $F = -k_B T \ln(Z)$, where Z is the partition function, Parisi and others followed Edwards and Anderson's strategy of using the identity

$$\ln Z = \lim_{m \to 0} \frac{Z^m - 1}{m}.$$

But in the spin-glass problem, m is the number of replicas, and researchers struggled with the implications of sending that value to zero.

Undeterred, Parisi forged ahead into the zero-dimensional space where the matrix containing the spin correlations $q_{\alpha\beta}$ resided. (Those correlations appear in the calculation of Z.) The analytic procedure he devised for dealing with such objects produced a stable and consistent description of the Sherrington–Kirkpatrick spin glass. Although the deeper meaning of the approach is still not fully understood, the solution's accuracy has since been rigorously validated.

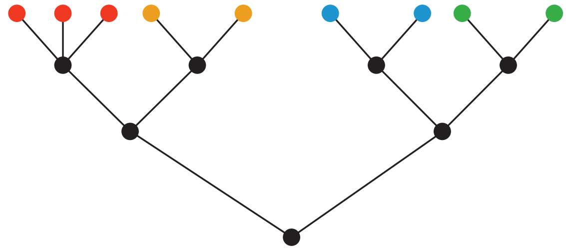


FIGURE 1. AN ULTRAMETRIC TREE describes the underlying organization of spin-glass states, represented by colored circles. The number of nodes separating two states indicates how similar they are. Any two states of the same color have the same amount of overlap, whereas a red and an orange state have less overlap—which is still more than either would have with a blue or a green state. (From P. W. Anderson, PHYSICS TODAY, July 1989, page 9.)

A phase transition is normally accompanied by a change in an order parameter—a value characterizing the system's degree of structure—usually from zero to nonzero. In a ferromagnet, that value is the magnetization, which is produced by spins aligning below T_c.

Parisi's order parameters $q_{\alpha\beta}$, of which there are infinitely many, instead measure the amount of overlap between two states. As expected, comparing a state with itself yields the largest overlap, $q_0 = q_{\alpha\alpha}$. But surprisingly, as Parisi and his collaborators showed in 1984, the overlaps between the rest of the seemingly unconnected states weren't random.[5] The most similar nonidentical states had overlap $q_{\alpha\beta} = q_1 < q_0$, followed by $q_2 < q_1$ for the next-closest states, and so on.

As illustrated in figure 1, the set of overlaps is ultrametric in structure—they satisfy the strong triangle inequality, $q_{\alpha\gamma} < \max(q_{\alpha\beta}, q_{\beta\gamma})$. The number of nodes between two states reflects their amount of overlap. For example, any two states of the same color in the figure have the same amount of overlap, q_1. A red and an orange state have the next-highest overlap, q_2, the same as a blue and a green state. The branches of the ultrametric tree appear as the system is lowered below T_c and divide further and further as the temperature decreases—a reflection of the free-energy landscape becoming increasingly rugged, with more and deeper minima to trap the system.

The spin-glass cornucopia

In simulations to investigate the zero-temperature entropy and other observables in spin glasses, Sherrington and Kirkpatrick encountered an obstacle: Once a spin glass settled into a local minimum, they couldn't efficiently explore its rugged free-energy landscape for other stable states. From that starting point, changing just a handful of spins always increased the system's free energy, leaving it on the steep slope surrounding the minimum. Parisi's solution made it possible to show that the energy barriers between states grew with N, the total number of spins. Any effort to systematically search the landscape of a many-spin system would be futile.

Still, real-world and simulated spin glasses manage to form in a wide variety of stable states. And heating a system into its disordered phase and then recooling it allows it to fall into different minima without getting stuck in a particular one. Repeating that enough times, Kirkpatrick realized, would map the landscape.

Using that insight he, Daniel Gelatt, and colleagues demonstrated that such simulated annealing could be applied to other problems of combinatorial optimization—that is, finding the best solution from many possibilities.[6] One example they considered was the "traveling salesman" problem, which involves finding the shortest route a salesman could take to visit N cities, each exactly once, and then return home (see figure 2a). The number of possible routes is on the order of $N!$, which quickly becomes intractable for even the most powerful computers. But even with thousands of cities, simulated annealing almost always effectively and efficiently finds the optimal solution, or at least a nearly optimal one.

The traveling salesman problem is just one of many combinatorial optimization problems, some of which are of significant practical importance. Air-travel routes, supply-chain networks, computer-chip design, and distribution logistics all require solving similarly complicated problems. Biological problems such as finding folded protein configurations (see figure 2b) also benefit from such search techniques.

In the early 1980s John Hopfield, then a neuroscientist at Caltech, was studying neural networks—neurons connected by synapses—and how their col-

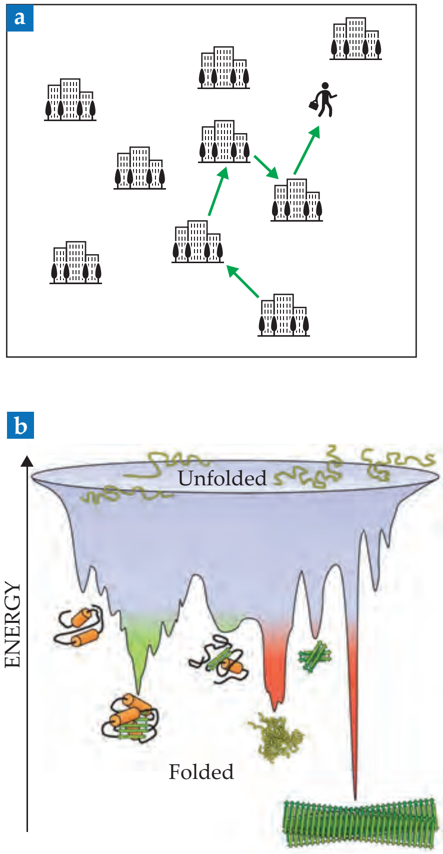


FIGURE 2. SPIN-GLASS CONCEPTS can be used to understand problems in a wide range of fields. **(a)** In the "traveling salesman" problem, one must find the shortest route connecting a set of cities. With more than a handful of cities, systematically solving the combinatorial optimization problem becomes prohibitively expensive on a computer. Simulated annealing, a computational technique inspired by spin glasses, significantly reduces that cost. **(b)** Proteins, like spin glasses, have complicated free-energy landscapes with many deep minima. Similar techniques can be used to study the dynamic evolution of both systems. (Adapted from F. U. Hartl, A. Bracher, M. Hayer-Hartl, *Nature* **475**, 324, 2011.)

lective behavior could result in phenomena such as memory formation. With neurons and synapses taking the place of spins and their interactions, respectively, Hopfield's network had a form that closely mirrored that in the Sherrington–Kirkpatrick model. The ultrametric structure of the solutions can also describe patterns in memory formation, in both people and computers.

Applying Parisi's formulation to Hopfield's model was a gateway for many theorists. "To those of us who knew about spin glasses, there was a great deal of similarity with what we'd been working on, and the opportunity to move in that direction," says Sherrington.

A natural question is whether spin-glass techniques can be applied to more ubiquitous structural glasses. Although some of the same ideas are relevant, namely disorder and frustration, the more familiar glasses have an added level of complication: Their disorder is not quenched. Once a spin glass's magnetic moments are in place, their interactions are fixed. But the components of other glassy systems—be they molecules, grains, or biological materials—move around as they approach a glassy state, so their interactions are always evolving.

"In a sense, the ruggedness of the free-energy landscape of the fragile glasses is even more complex than for the spin glasses," says Cugliandolo. Still, they share many properties, and some of the same tools can be applied. Parisi and his collaborators have used the replica method to exactly solve structural-glass problems in infinite spatial dimensions. Their mean-field solutions show that such glasses have a so-called Gardner transition, a nonequilibrium phase transition regarding the structure of free-energy basins, that was first identified in spin glasses by Elizabeth Gardner in 1985.

The cavity method, an iterative process developed by Parisi, Marc Mézard, and Miguel Virasoro as a mathematical alternative to the replica method, has been applied algorithmically to problems in constraint satisfaction, information theory, and other areas. Indeed, a wide range of fields that deal with complex interaction networks have benefitted from spin-glass concepts and methods, including economics, ecology, and evolution.[7] And the list continues to grow.

A universal scientist

Parisi's work on spin glasses and replica-symmetry breaking has contributed to the understanding of complex and disordered systems in a range of fields. But, says Cugliandolo, "He has worked in so many different branches of physics and made important contributions to all of them, so I think this was also very important in the award."

With researchers pursuing more siloed and specialized problems, it's increasingly rare to hear about physicists making such broad contributions. But as a field theorist, Parisi has found wide-ranging applications for his skills. "He's mentored a huge number of people who have gone all over the world and made great contributions in many different fields," says Andrea Liu, a professor at the University of Pennsylvania in Philadelphia. "He's very encouraging to people and works with students really well."

Parisi first made a name for himself in quantum chromodynamics, the study of interactions between quarks and gluons. He and Guido Altarelli developed a quantitative description of inelastic scattering among quarks and gluons; their 1977 paper on that work is Parisi's most cited.[8]

Perhaps the clearest connection between Parisi's contributions and the other half of the Nobel Prize recognizing climate modeling (see page 14 of this issue) is his work on stochastic resonance and turbulence.[9] He and his collaborators found that in nonlinear systems, just the right amount of noise combined with periodic forcing can generate a positive feedback loop and unexpected behavior. In particular, solar radiation and noise from oceanic and atmospheric dynamics cause large, seemingly paradoxical periodic temperature variations on the order of 10 °C.

Parisi's willingness to tackle complex problems has led him to make many fundamental contributions, and his Nobel Prize celebrates that type of work. "These are ideas that come from physicists working in the way that physicists do," says Sherrington. "If he hadn't been working on very basic things, asking these very interesting questions, you wouldn't have gotten to the other places that have been so useful."

Christine Middleton

References

1. G. Parisi, *Phys. Rev. Lett.* **43**, 1754 (1979).
2. S. F. Edwards, P. W. Anderson, *J. Phys. F: Met. Phys.* **5**, 965 (1975).
3. D. Sherrington, S. Kirkpatrick, *Phys. Rev. Lett.* **35**, 1792 (1975).
4. J. R. L. de Almeida, D. J. Thouless, *J. Phys. A* **11**, 983 (1978).
5. M. Mézard et al., *Phys. Rev. Lett.* **52**, 1156 (1984).
6. S. Kirkpatrick, C. D. Gelatt Jr, M. P. Vecchi, *Science* **220**, 671 (1983).
7. D. L. Stein, C. M. Newman, *Spin Glasses and Complexity*, Princeton U. Press (2013).
8. G. Altarelli, G. Parisi, *Nucl. Phys. B* **126**, 298 (1977).
9. R. Benzi et al., *Tellus* **34**, 10 (1982); R. Benzi et al., *J. Phys. A* **17**, 3521 (1984).

Ten billion years ago, galaxies were already running out of gas

A look into the distant reaches of the universe offers clues about when, why, and how star formation shuts off.

Our sun, at roughly 4.5 billion years old, is a relative latecomer to the universe. Star formation had its heyday more than 10 billion years ago, 3.5 billion years after the Big Bang. It's been declining exponentially since then. (See the article by Jeremiah Ostriker and Thorsten Naab, PHYSICS TODAY, August 2012, page 43.)

But not all galaxies were invited to the party. Even when the rate of star formation was at its peak, half of the most massive galaxies were creating almost no new stars. Quiescent galaxies have only increased in number since then, because once a galaxy's star formation shuts off, it rarely turns back on.

Astronomers don't know why some galaxies stop forming new stars, why the shutdown seems to be so sudden and permanent, or why it happens for some galaxies so early. Galaxies distant enough to provide a window into what was happening in the universe so long ago are often too faint to study in the requisite detail.

Now the REQUIEM collaboration (short for "Resolving Quiescent Magnified Galaxies"), co-led by Katherine Whitaker of the University of Massachusetts Amherst and Sune Toft of the Niels Bohr Institute in Copenhagen, has identified a new piece of the puzzle.[1] In a study of half a dozen quiescent galaxies, chosen because their Earthbound light is boosted by gravitational lensing, Whitaker and colleagues found that the galaxies had literally run out of gas: They stopped forming stars because they lacked the necessary cold hydrogen gas to make them.

The result is far from the final word on the matter. It doesn't explain where the gas went—was it simply used up and never replenished, expelled from the galaxy somehow, or heated up into a form that's incapable of condensing into stars? Nor does it prove that all quiescent galaxies are quiescent for the same reason.

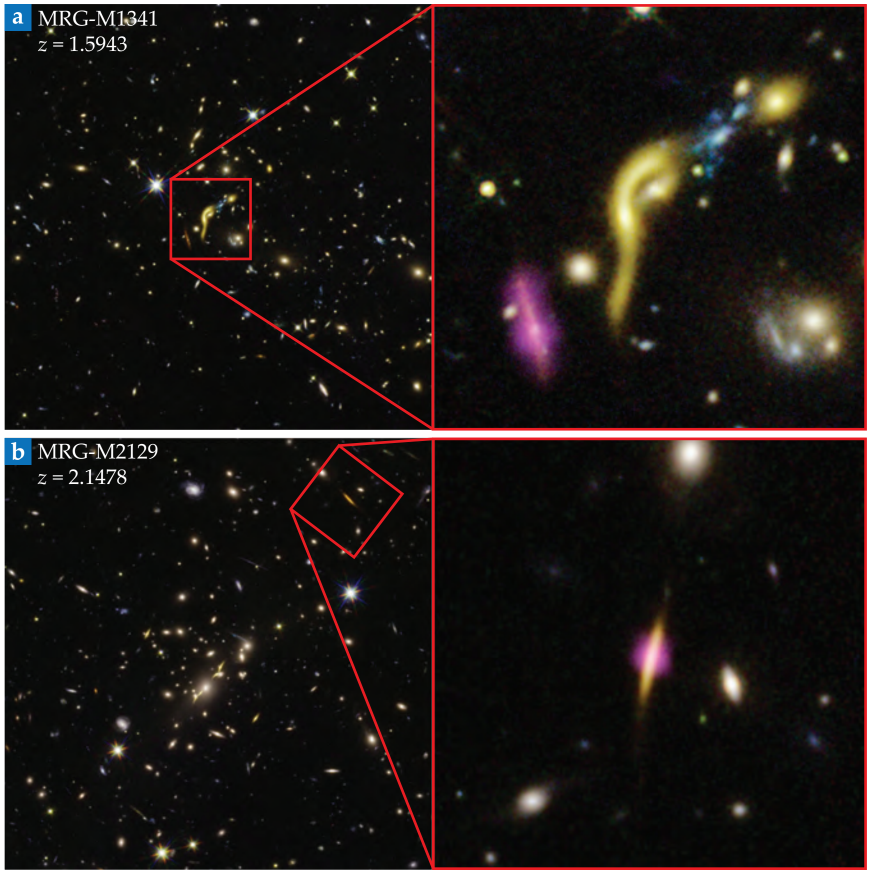


FIGURE 1. GRAVITATIONAL LENSING allows researchers to hunt for dust—a proxy for cold hydrogen gas, the stuff of star formation—in distant galaxies at high redshift z. Here, the dust signal (purple) is superposed on images from the *Hubble Space Telescope*. **(a)** Several highly lensed galaxies, including the one at the center of the right panel, produce no detectable dust signal at all. **(b)** Others show a weak dust signal that, mysteriously, appears to emanate from only part of the galaxy. (Images courtesy of NASA, the European Space Agency, Katherine Whitaker, and Joseph DePasquale.)

But the result does suggest that at least one mechanism is capable of depleting a galaxy's star-forming gas extremely early in the universe's history—at a time when most other galaxies had plenty.

Counting dead galaxies

Stars in galaxies other than our own are too distant for telescopes to distinguish, let alone to observe their creation. So how do astronomers know which galaxies are forming new stars and which aren't?

The key is that not all stars are created equal. The heaviest main-sequence stars, with masses more than 10 times that of the Sun, are bright, hot, and blue, and they burn themselves out in as little as 3 million years. At the other end of the continuum, stars less than a tenth of a solar mass are cooler and redder, with expected lifetimes of hundreds of billions of years.

A galaxy's color is thus a record of the kinds of stars it contains. If it emits a lot of blue light, it must be actively making new stars to replenish the short-lived blue stars that are burning themselves out. If blue light is absent, on the other hand, the galaxy is no longer making blue stars, or indeed any stars: It's quiescent—or more colloquially, "red and dead."

For extremely distant galaxies, the measurement is a bit more complicated, because a galaxy's light can be reddened either by the cessation of star formation or by the redshift caused by the expansion of the universe. Those two phenomena have different effects on the shape of

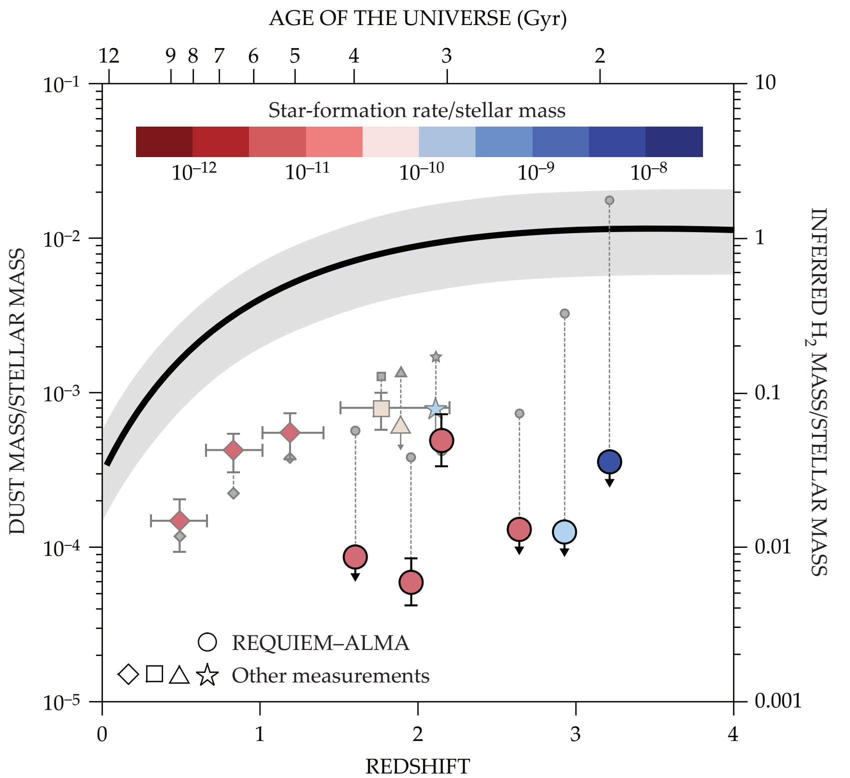


FIGURE 2. SURPRISINGLY LITTLE dust and hydrogen gas is observed in six quiescent REQUIEM target galaxies (colored circles). The amount is low compared with prior measurements[3] of other quiescent galaxies, the expected average for massive galaxies (solid line), and even the expectation based on the galaxies' own star-formation rates (gray symbols). The colors represent the fraction of each galaxy's mass forged into new stars per year. The most strongly redshifted galaxy (dark blue) has a high star-formation rate as averaged over the past 100 million years; it's included in the sample because its star formation has rapidly shut down since then. (Adapted from ref. 1.)

the spectrum, and astronomers are adept at disentangling them to separately measure a galaxy's redshift z and star-formation rate.

To study a distant galaxy in detail, however, the faint, redshifted light often doesn't suffice. That's where REQUIEM comes in, with the idea to exploit gravitational lensing. If a distant galaxy lies directly behind a massive foreground object, the object's gravitational heft can exert such a pull on the galaxy's light that more of the light is bent toward Earth.

That configuration is rare in absolute terms, but out of the billions of galaxies in the universe, some are aligned just right. From a decade of combing through galaxy survey data, the REQUIEM researchers identified 10 suitable target galaxies.[2] All are magnified by lensing by up to a factor of 30. All are distant, with redshifts between 1.6 and 3.2, meaning that their light dates from between 9.5 billion and 11.5 billion years ago. And they're all quiescent: While most are fully red and dead, a few contain some stars blue enough to be less than 100 million years old but have rapidly stopped forming new stars since then.

Gas and dust

The present work uses the Atacama Large Millimeter/Submillimeter Array (ALMA) in Chile to study six of the REQUIEM target galaxies—all that the facility's time would accommodate so far. The ALMA signal, at a wavelength of 1.3 mm, comes not from the galaxies' stars but from their interstellar dust, a proxy for H_2 gas.

Measuring the galaxies' H_2 content directly isn't an option because H_2 molecules themselves are essentially invisible. They're symmetric and lack electric dipole moments, so they don't absorb or emit radiation when they rotate and vibrate. They're visible to telescopes only when they're hot enough to excite into higher electronic states, not when they're cold enough to condense into regions dense enough to become stars.

To observe cold H_2 gas, astronomers need to look for some other substance that's found with it. One option is carbon monoxide, an asymmetric molecule that does have a dipole moment and thus a cold spectral signature. But observing CO is expensive, because it requires greater spectral resolution than most telescope images provide.

The other option is to look for dust. Dust grains grow when gas molecules stick together, explains Whitaker, and "in the local universe, we observe that where you have gas, you have dust." Typical local galaxies contain 100 times as much H_2 as they do dust, and the REQUIEM researchers assumed that their target galaxies do as well. But they don't know for sure—precisely because the dust content has never been measured before in galaxies at such high redshifts. "That's a big assumption that goes into this work," says Whitaker.

The data were collected during 2018 and 2019. (ALMA suspended its operations in March 2020 due to the COVID-19 pandemic.) As each new batch of data arrived, the researchers grew more surprised: The dust signal was almost entirely absent. Four of the six galaxies, including the one shown in figure 1a, showed no observable dust at all; the ALMA signal, shown in purple in the figure, emanated only from other objects in the field of view. The best the researchers could obtain was an upper bound on the dust content.

The other two galaxies, including the one in figure 1b, exhibited weak dust signals. But they presented another mystery: The dust wasn't shaped like the rest of the galaxy. Gravitational lensing caused each galaxy to appear elongated, and the dust signal is confined to only part of it. "What's going on there?" asks Whitaker. "We still don't understand what that means."

A mysterious absence

Overall, as shown in figure 2, the six REQUIEM galaxies (colored circles) had a median dust mass fraction of less than 0.01%, from which the researchers inferred a median H_2 mass fraction of less than 1%. That's a surprisingly small number by several standards. It's less than what was observed in other studies (the other colored symbols in figure 2) that focused on a slightly lower redshift range.[3] It's less than the average amount of gas that must have been present in massive galaxies to power their star formation (the solid line). It's even, in most cases, less than the amount of gas that would be expected based on the galaxies' own rates of star formation (the small gray circles).

"But I think the most important thing," says Whitaker, "is that these are null detections. We're still not going deep enough to really constrain how little cold gas and dust are there." Among the team's next steps is to keep collecting more data with ALMA—now back up and running after its pandemic shutdown—to tighten their upper bounds on the galaxies' dust content.

Because the REQUIEM study included just six galaxies, the researchers can't tell whether the lack of dust and gas is common to quiescent galaxies in general or an unusual feature of the few they happened to sample. Small statistics are an inherent limitation of REQUIEM's approach because strong gravitational lensing of galaxies is rare. "We may eventually find a few more galaxies we can study this way," says Whitaker, "but I doubt we'll ever get to order 100."

Even so, the complete absence of cold H_2 gas in even a few galaxies so early in the universe's history is difficult to reconcile with current understanding of cosmology and galaxy evolution. Ten billion years ago, the universe was much smaller than it is now, and the concentration of intergalactic gas was much higher. Furthermore, intergalactic space was (and still is) permeated by a web of dark-matter filaments that gravitationally attract gas and channel it toward galaxies.

Theorists do have some ideas for how gas could be depleted, expelled, or overheated in early galaxies. Many of the proposed mechanisms involve active galactic nuclei (AGNs): supermassive black holes at galactic centers whose rapid accretion of matter may power the heating or expulsion of a galaxy's gas. As the REQUIEM researchers continue to gather data on their galaxies with more instruments and in more wavelength ranges, much of their effort will focus on looking for the observational signatures of AGNs.

The soon-to-be-launched *James Webb Space Telescope*—already scheduled to gather data on the galaxy in figure 1a—could help considerably. The telescope is optimized to observe in the near-IR, a frequency range where spectral features from redshifted galaxies lie but that's difficult to observe with other existing telescopes. Furthermore, it can image in the near-IR with simultaneously high spatial and spectral resolution. So, for example, it would be able to detect gas flowing outward from an AGN at the center of a galaxy. Says Whitaker, "It really is the perfect telescope for understanding the chemical history of these galaxies."

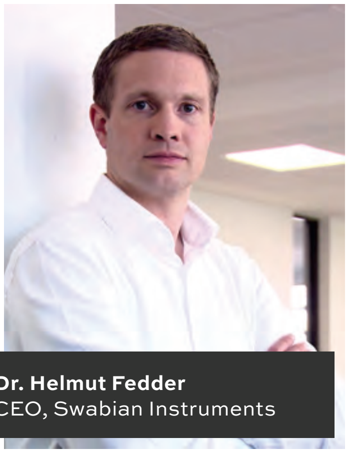
Johanna Miller

References

1. K. E. Whitaker et al., *Nature* **597**, 485 (2021).
2. See, for example, H. Ebeling et al., *Astrophys. J. Lett.* **852**, L7 (2018); A. B. Newman et al., *Astrophys. J.* **862**, 125 (2018); S. Toft et al., *Nature* **546**, 510 (2017).
3. J. Zavala et al., *Astrophys. J.* **887**, 183 (2019); J. Caliendo et al., *Astrophys. J. Lett.* **910**, L7 (2021); R. Gobat et al., *Nat. Astron.* **2**, 239 (2018); G. Magdis et al., *Astron. Astrophys.* **647**, A33 (2021).

PT

Swabian Instruments paves the way to optical quantum computing



Dr. Helmut Fedder
CEO, Swabian Instruments

Swabian Instruments is a test and measurement company headquartered in Stuttgart, Germany. We strive to make data acquisition and signal generation as easy and intuitive as possible.

Swabian Instruments was founded by a team of physicists and engineers led by Helmut Fedder. The researchers needed an outstanding acquisition system for their own experiments at the University of Stuttgart. Ultimately, they decided to develop the dream device themselves. The resulting Time Tagger series proved to be so successful in the scientific community that in 2016 the developers decided to make the time-tagging technology available to researchers worldwide.

Today, Swabian Instruments is a rapidly growing company of more than 20 experts. A crucial part of our success is our collaborations with users in academia and industry. This strong customer focus enriched our technology, gave us an invaluable insight into the most advanced scientific labs, and rapidly made the Time Tagger one of the market leaders in single-photon counting.

Swabian Instruments established its expertise just in time to contribute to one of the hottest strategic technologies: optical quantum computing. Development in this area demands high timing resolution, simultaneous handling of multiple coincidence groups, modern and flexible acquisition control, and on-the-fly data processing. Swabian Instruments has all that to offer.

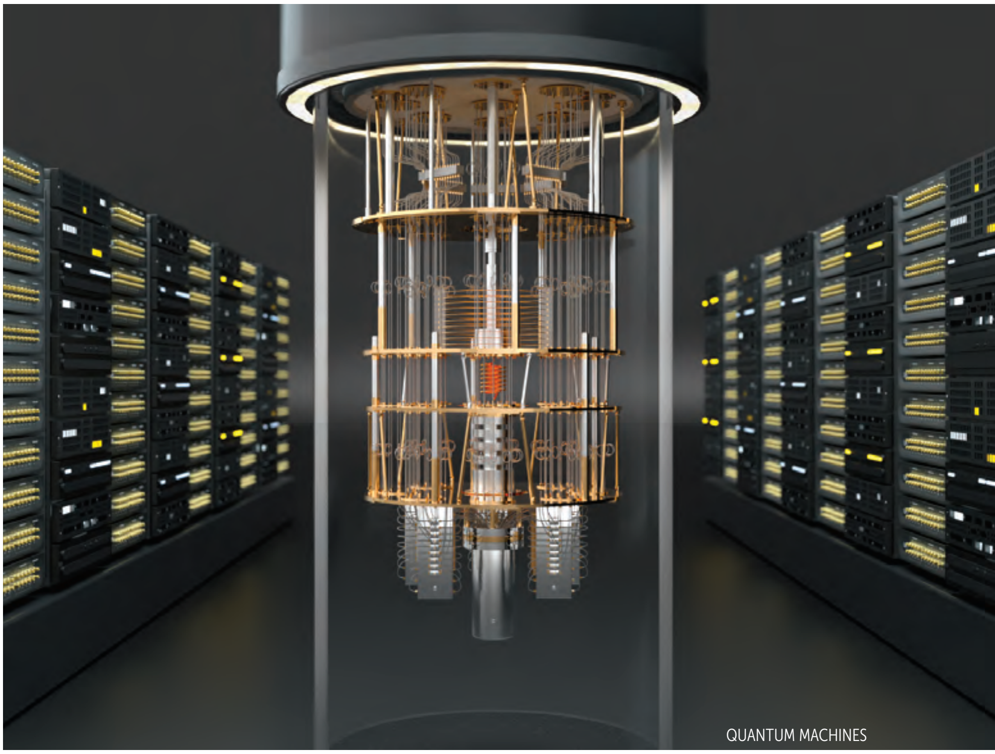
Therefore, our entry into optical quantum computing happened naturally, through numerous seminal projects with our early adopters. Early on, we joined efforts with the group of Ian Walmsley (now at Imperial College London), who led the Networked Quantum Information Technologies Hub in Oxford. Time Tagger came in handy for his postdoc, Benjamin Brecht (currently at Paderborn University), who developed a patented quantum memory protocol for the storage and retrieval of heralded single photons. Jianwei Wang (now at Peking University) employed our multichannel system while working at the University of Bristol on a new generation of integrated quantum optics toolbox with on-chip handling of light's quantum states.

Seeing quantum computing through the eyes of Swabian's users helped us tailor our electronics and software to their diverse needs. Multiple successful collaborations proved our competence and ability to quickly deliver results in this dynamic field. Funded research and development projects within large academic and industrial consortia have followed. Swabian Instruments, which has unique expertise in event timing, now contributes to optical quantum computing shoulder to shoulder with such technology leaders as Fraunhofer IOF, Ulm University, Jülich Research Center, the University of Jena, Menlo Systems GmbH, the Technical University of Munich, Karlsruhe Institute of Technology, and Fraunhofer IPMS.

Swabian Instruments' user community will soon benefit from our latest development, Time Tagger X, which features an unprecedented timing resolution of less than 2 ps, 18 channels on board, and fast direct access to the time tags. Because of our focus on innovation and emerging applications, the world's technology experts rely on our scalable state-of-the art instruments as they help shape the future of quantum computing.

Israel has become a powerhouse in quantum technologies

A supportive government, available capital, and world-class academic institutions are some of the factors behind the nation's quantum ascendance.



A SUPERCONDUCTING QUBIT is, in this rendering, controlled by quantum orchestration hardware and software made by the Israeli company Quantum Machines.

Ron Folman, a quantum physicist, was taken aback when a delegation of top officials from the US Defense Advanced Research Projects Agency paid an unexpected visit to his lab at Ben-Gurion University of the Negev (BGU) about 15 years ago. "I asked them, 'Why are you here? You have a thousand times more money, scientists, and space,'" Folman recalls. "The head of the delegation answered, 'I heard what you guys are doing, and I wanted to see with my own eyes this Israeli chutzpah.'"

To Folman, the quote neatly characterizes Israel's outsize global footprint in quantum science and technology.

It's no accident that Israel is punching above its weight in quantum fields. Despite the turmoil that has roiled Israeli politics in recent years, the Knesset committed 1.25 billion shekels ($400 million) to a five-year National Quantum Initiative, which kicked off in late 2019. Tal David, an experimental physicist who heads the initiative, says it gained a boost from Israel's economic stimulus program during the COVID-19 pandemic. The program includes $60 million to build the country's first quantum computer, which is expected to consist of 30–40 qubits. Assembly is due to get underway early next year.

"We don't aspire at least in the next few years to beat IBM or Google," says David. "We need to first form a basis for an ecosystem in Israel and that will be the purpose of the project. Maybe in a few years, we'll be able to jump into the deep end of the pond and compete with the big ones."

Today 60% or more of the initiative's funds is spent in academia, says David. Israel consistently ranks high alongside leading European nations by such measures as citations and the number of grants in quantum fields awarded by the European Research Council, of which it is an associate member.

But the initiative's center of mass is beginning to shift to industry, David says, which should further stimulate the explosive growth of the last two years, during which the number of Israeli firms working in quantum tech surged from 5 to 30. The new entries span the gamut of hardware and software for defense and civilian applications. "It's a small ecosystem, but it's developing quite rapidly, and it's quite diverse," he says, adding that "people are trying not to step on each other's toes."

Attracting $75 million in private investment to date, three-year-old Quantum Machines ranks among the top quantum startups in the world by that measure. Mellanox Technologies, which entered the quantum communications business two years ago, was acquired by Nvidia in 2020. Most of the tech giants, including Microsoft, Google, Amazon, Intel, and others, have R&D centers in Israel that serve as spawning grounds for new quantum companies.

Adopting the nickname "Start-Up Nation" from the eponymous best-selling 2009 book, Israel has an existing high-tech innovation ecosystem that includes a government eager to back inventors. That support in turn attracts investment from the nation's substantial venture capital community. "When they see the government is pushing hard, it's easier for them to take the risk," says David. "At some point we are sharing risk with the private sector."

When it comes to high tech, "basically anyone who comes with an idea and is able to prove the concept probably will get some kind of government grant to kick it off," says Shlomi Cohen, CEO of QuantLR, a quantum encryption startup. "At any moment in Israel, there are more than 6500 startups. I think we are second only to the US." (See PHYSICS TODAY, October 2021, page 42.)

The Israeli army, air force, and intelligence community are the "backbone of the [quantum] industry," Cohen says. The military is both a consumer of quantum technology and a source of technical talent for startups. The government-owned defense contractor Rafael Advanced Defense Systems has technology-sharing agreements with academic institutions, including BGU and the Weizmann Institute of Science. Israel's defense industry began taking an interest in quantum sensing about a decade ago, says Nir Davidson, a Weizmann physicist. The national initiative has allocated $40 million to a quantum-sensing consortium composed of five companies and eight academic groups, including Folman's and Davidson's labs.

A similar consortium is organized around quantum cryptography, where David says activity is "moving much faster than we anticipated" from universities to industry, and into applications in large organizations and the government. A third consortium aims to enlarge Israel's presence in quantum computing, where, he says, "we have a very good start, but it's small."

Well positioned

Davidson is developing a cold-atom gravimeter that will be useful both as an accelerometer in inertial navigation systems and in basic science. LightSolver, a quantum-inspired technology, emerged from Davidson's lab. While not technically a quantum computer, the desktop device that is under development will use a coupled laser array to perform computations.

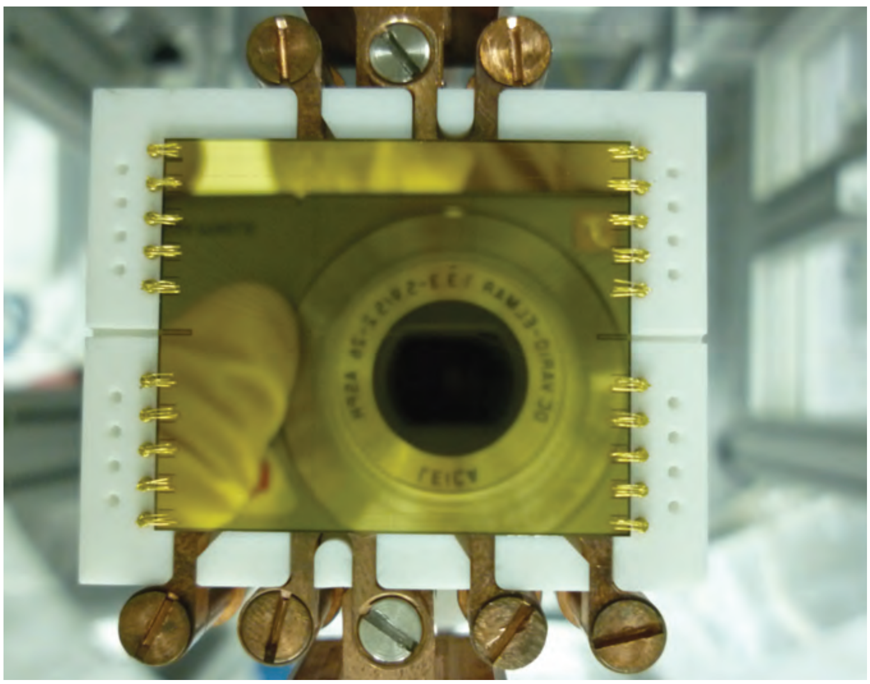
Davidson says the jury is still out on whether quantum computation will come into widespread use in the next decade, but "if it does it will be a game changer, and people will invest huge sums." Well positioned to benefit in that event is Quantum Machines, founded by three physicists who earned their PhDs in the lab of Weizmann physicist Moty Heiblum. The company's quantum orchestration platforms communicate complex software to quantum processors and perform quantum error correction. Notably, the devices will work with whatever type of quantum computing platform becomes dominant. (See "What's under the hood of a quantum computer?," PHYSICS TODAY online, 5 March 2021.)

Itamar Sivan, CEO and cofounder of Quantum Machines, says he has no doubt that quantum computing will take off. "The question is when." He attributes much of his 70-employee company's success to the ready availability of funding. "And there are great engineers and amazing talent in Israel. We can find people here who are both experts in quantum but also have some engineering background." Many of them have come from the defense sector.

Folman's lab has already delivered two quantum prototypes to industry: a compact, robust cold-atom atomic clock and a sensitive magnetic atomic sensor. Amir Waxman left Folman's lab seven years ago to work for AccuBeat, an Israeli manufacturer of rubidium clocks. He continues working with BGU researchers to commercialize the new clock technology he helped to develop there.

Whereas today's commercial atomic clocks are accurate to within one second in 2000 years, the new type will lose just one second in 1 million years, Waxman says. That level of accuracy is important for myriad applications including radar, financial transactions, phase synchronization in electricity transmission, and inertial navigation. For example, a nanosecond can produce errors of several meters in distance as measured with radar, and it is critical to determine the exact sequence of trades in financial instruments, which often occur microseconds apart, he explains. The lab is now developing a miniaturized ytterbium optical-frequency atomic clock that will be five orders of magnitude more accurate than today's clocks.

Folman's lab is working on atom chips—cold-atom devices that include miniature particle sources, vacuum pumps, lasers, and sources of electromagnetic fields to prepare, manipulate, and measure individual atoms. The chips also contain electronics and fiber optics for readout. Just as the miniaturization of the transistor led to the electronic revolution with devices such as laptops and cell phones, atom chips could enable miniaturization of quantum technology, allowing greater accuracy and complexity and lower power consumption, Folman says.


YAIR MARGALIT

AN ATOM CHIP built at the Ben-Gurion University of the Negev nanofabrication facility. The chip is part of a sensor known as a Stern–Gerlach atom interferometer.

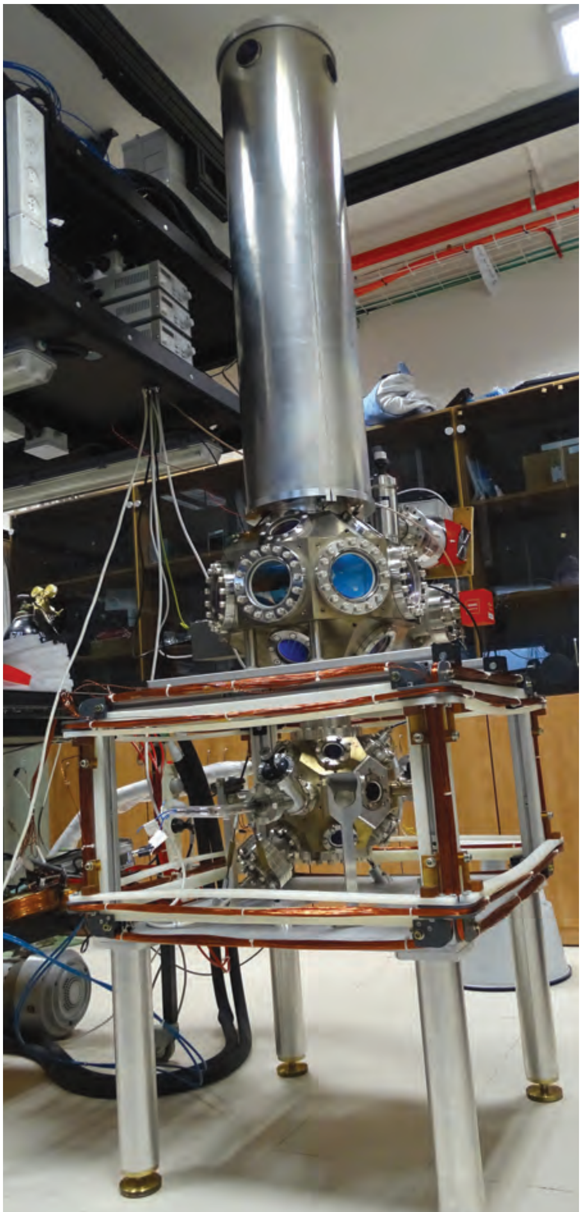
Quantum sensors may map the human brain, enable quantum simulators for drug development, and allow long-range quantum communication. Quantum gravitational sensors could produce subsurface images that predict earthquakes and volcanic eruptions. For the military, where stealth technology is in danger of being overcome, "no one knows how to counteract gravitational signals," Folman notes.

To accommodate the novel designs, geometries, and materials required for atom chips, BGU has built what Folman says was likely the world's first fabrication facility dedicated to atom-chip R&D. It has orders from R&D centers in Germany, the UK, the Netherlands, Italy, and the US. Companies such as ColdQuanta in Colorado and IonQ in Maryland are already producing such chips for commercial use. (See PHYSICS TODAY, November 2020, page 22.)

By teaching Israelis how to team up to accomplish tasks, mandatory service in the Israel Defense Forces also contributes to the nation's tech success, says Folman. "They learn to complement and trust each other." But that's not enough, he adds. "At the end of the day, you need innovation and ingenuity. People in Israel are very passionate about innovation."

An ecosystem

Serendipity played a role in both Raicol Crystals' and Tabor Electronics' entry into quantum fields. The two Israeli companies had long-established business lines when quantum technologies began to take off about five years ago. Founded in 1995, Raicol, which manufactures crystals for the laser and optics industries, discovered that physicists were using its nonlinear periodically poled potassium titanyl phosphate crystals in their research to produce entangled photons, through a process called spontaneous parametric down-conversion. Raicol responded to that nascent market by tailoring its crystals for quantum computing, communications, and sensing applications, says Ori Levin, vice president of product management and strategy.



CHEN AVINADAV AND DIMITRY YANKELEV

A QUANTUM GRAVIMETER setup at the Weizmann Institute of Science. The device was built in collaboration between the Weizmann Institute and Rafael Advanced Defense Systems, the Israeli defense contractor.

A decade ago, 50-year-old Tabor Electronics learned that some of its physicist customers were using its standard signal generator products in their quantum research. Earlier this year, the company introduced a dedicated arbitrary waveform transceiver designed specifically for quantum computing, communication, and sensing, says Mark Elo, the company's vice president and general manager of the Americas. The Israel Innovation Authority provided funding to support development of the new instrument.

The two-year-old startup QuantLR expects to start delivering systems for quantum photon data encryption in about a year, says Cohen. "We are using superposition of photons to send the key encryption," he says. "No one can penetrate a photon. That's the basic idea of our approach."

QuantLR's hardware and software will secure data transmitted through fiber-optic cable over distances up to 100 km, beyond which cooling systems would be required. Most fiber-optic communications between data centers, financial institutions, utilities, infrastructure, and 5G base stations occur within that range, he says. Intrusions will create noise that can be detected, he says, and data flows won't be slowed because the quantum key will travel over two dedicated channels in the fiber cable, rather than being attached to the data itself.

The company is "in intensive discussions" with data center operators including Amazon, Microsoft, and Telefónica and with 5G companies including Cisco, Nokia, Ericsson, and Huawei.

Looking ahead

David believes there is more to be done. Israel's number of principal investigators "at the core of quantum," about 125, is low even for a small country, he says. By comparison, about 300 principal investigators worked at the basic science level in Israel's 10-year national nanotechnology initiative, which ended in 2016. "One of the biggest challenges we have is to enhance the community—academia as well as industry—while maintaining the high degree of scientific and technological excellence," he says.

Davidson says that despite its success with European Research Council grants, "where the only criterion is scientific excellence," Israel hasn't done as well with international collaborations that are funded by the European Union's 10-year Quantum Flagship program, which began in 2018.

Other disciplines must get involved to fully exploit quantum's potential. "Even if you build the heart of a quantum device, there are many layers and technologies needed to be able to work," says Folman. But materials science, engineering, chemistry, and even biomedicine have yet to team up with physics and computer science, either in Israel or worldwide. "These disciplines today do very little connected to quantum technology. It should be an interdisciplinary effort, and eventually it will be."

Shai Lev, head of business development and partnerships at the quantum algorithm design company Classiq, took on a pro bono role as cofounder of Qubit, a grassroots quantum community effort meant to attract disciplines outside of physics and computer science into the quantum world. While most of the periodic seminars the group organizes are highly technical, some are also devoted to the financial aspects of starting a business. Shir Peri Lichtig, a product operations manager at Soluto and another Qubit co-organizer, says the sessions soared in popularity and became international after the pandemic forced them into virtual mode. Since then, such quantum oracles as Terry Rudolph from Imperial College London; John Martinis from University of California, Santa Barbara; and Chad Rigetti, CEO of Rigetti Computing, have contributed talks.

Paradoxically, another reason for Israel's big presence in tech may be its tiny size. "The maximum distance between universities here is like 200 kilometers," says Lev. "People know each other, and people are always looking for the next thing, and around them you have a vibrant community."

David Kramer

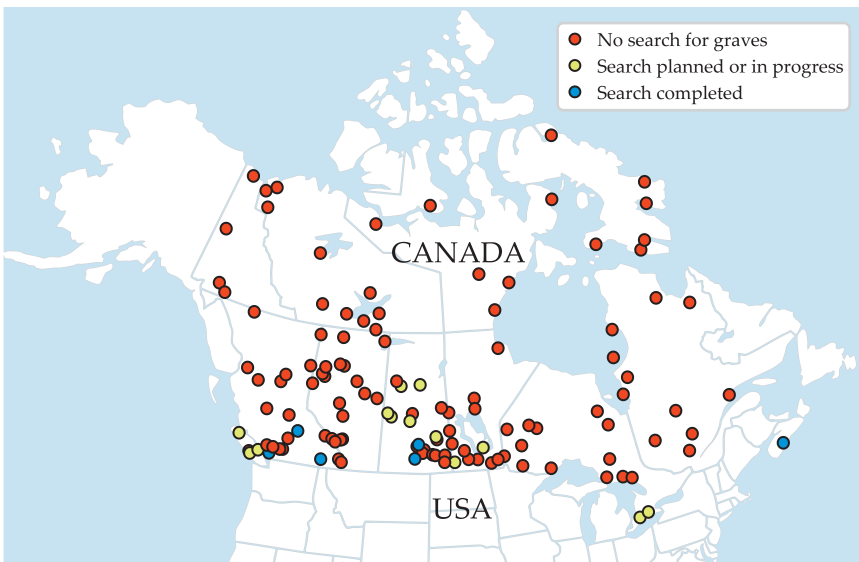
Subsurface imaging shows scale of the tragedy of Indigenous children

Archaeologists and geophysicists in Canada are trying to pinpoint the signatures of unmarked burials by using geophysical methods in varied terrains.

Few applications of radar pulses to locate objects underground have been as emotionally laden as identifying the graves of Indigenous children in Canada. The more than 200 burials located at the former Kamloops Indian Residential School in southwest British Columbia drew worldwide attention this past summer.

As part of the national government's attempt to assimilate Indigenous children into the dominant colonial culture, children as young as six years of age were forcibly taken from their families and sent to that school and many others across the country (see map) that were operated primarily by the Roman Catholic Church between 1890 and 1969. The US government enacted similar acculturation policies at American Indian Residential Schools in the same decades. Many children never returned home.

Ground-penetrating radar (GPR) has been used for several decades in engineering, geoforensics, and archaeology. (See PHYSICS TODAY, March 2014, page 24.) The Kamloops finding has forged a new interdisciplinary partnership to search throughout North America for other child



ARCHAEOLOGICAL SURVEYS that rely on ground-penetrating radar have identified unmarked children's graves at several former residential schools for Indigenous children in Canada. (Adapted from data from the National Centre for Truth and Reconciliation, University of Manitoba, Winnipeg, Canada.)

burial locations without disturbing them. Kisha Supernant, a Métis Indigenous woman and an archaeologist at the University of Alberta, uses maps and spatial data to explore how people in the past interacted with landscapes—for example, comparing the winter mobility patterns of different groups. Liam Wadsworth joined her research group in 2018 as a graduate student. He brought a geophysics background, training in GPR techniques, and a strong motivation to help Indigenous communities find unmarked graves. "Our goal is to make it possible for descendants to appropriately mourn the loss of their ancestral children," says Supernant.

Hyperbolic hints

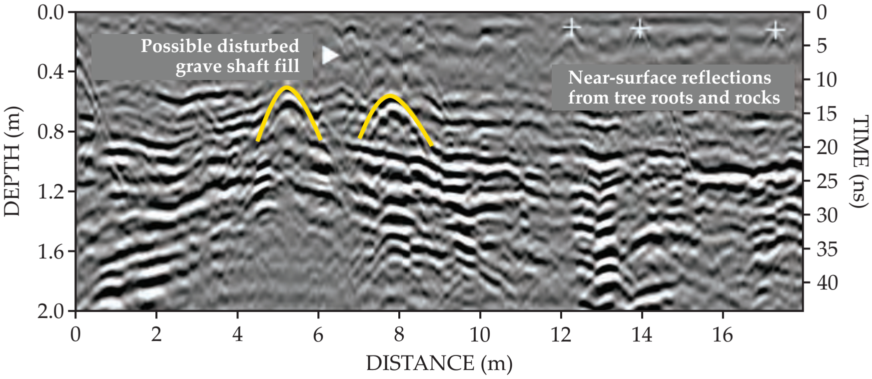
In GPR technology, an antenna transmits high-frequency radio pulses into the ground and records the return signal and the round-trip travel time. Travel times are sensitive to subsurface discontinuities that reflect, refract, or otherwise scatter the signal. Continuous profiles of subsurface reflections, called radargrams, provide an image of those discontinuities as the antennas are moved along the ground surface on lawnmower-like carts. A change in the return time of the transmitted radio waves could indicate an object, a gap, or a variation in ground composition. GPR today serves as a workhorse for mapping near-surface geological anomalies and buried utilities.

In a 1986 paper originally presented at an annual meeting of the Society of Exploration Geophysicists, Chris Vaughan discussed one of the first attempts to use GPR to search for graves. He examined a high-resolution survey of a 16th-century Basque whaling village on Canada's Labrador coast for signs of historical graves. The survey detected strong background signals that, through exhumation, were found to be associated with the burial sites. Later work by Kenneth Kvamme, an emeritus professor of anthropology at the University of Arkansas, described how that nondestructive tool could be used not just to scan individual burial sites but also to map spatial patterns over the many tens of hectares that characterize most archaeological excavations. That capability to map patterns was exactly what was needed to locate the many unmarked and undocumented graves of Canada's Indigenous children.

Those early results were promising, but their strictly empirical nature made it difficult to confidently apply GPR to areas where confirmatory exhumation was infeasible or culturally prohibited. In 2004 Lawrence Conyers, an anthropologist at the University of Denver, unraveled much of the physics behind the strong signal associated with graves. He developed methods that use the size, shape, and depth of the return signal to identify coffins and the associated vertical shaft features.

In addition to the depth of burials, important features for GPR are soil conditions, particularly the amount of clay that can attenuate the GPR wave, and the age of the grave. Older graves may hold remains in advanced stages of decomposition. The decomposed matter may make the distinction between the grave itself and the surrounding soil conditions difficult to discern.

Conyers found that GPR surveys using signals in the range of 300–500 MHz produce a good balance between penetration depth and resolution, both of which are needed to identify graves. In particular, he found that the disturbed



CONTINUOUS PROFILES of subsurface reflections, called radargrams, reveal subsurface features at the Cowper Lake Burial Ground on traditional territory of the Chipewyan Prairie First Nation, a designated group of Indigenous peoples in Alberta, Canada. The yellow curves indicate the characteristic "upside-down U" features that result from wooden coffin or shroud burials. The disturbed soil above those reflections may represent grave shafts. Surface reflections from roots and rocks illustrate the complexity of interpreting a ground-penetrating radar profile. (Adapted from W. T. D. Wadsworth et al., *Adv. Archaeol. Pract.* **9**, 202, 2021.)

soil of a narrow grave shaft creates a distinctive "upside-down U" anomaly in the return signal. The radar pulse reflects off the discontinuity at the top of the coffin or the bottom of a hollowed-out chamber. Steep-sided pits concentrate the reflected waves above them, creating the characteristic convex shape (see the image above). "If you get a hyperbola with a wave reverberating beneath it in a graveyard, there's hardly any explanation besides a grave," says Alastair Ruffell, a geophysicist at Queen's University Belfast who specializes in GPR for forensic research.

A disturbing problem

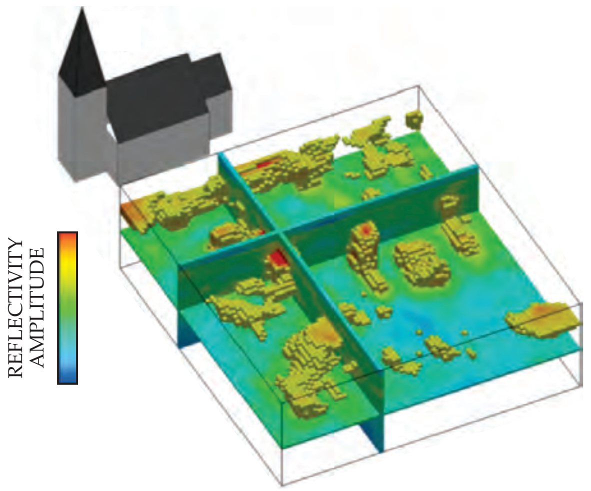
In 2012 the University of Manitoba's Katherine Nichols, then writing her master's thesis in forensic anthropology, heard a rumor that the list of 11 students on a cemetery marker near the Brandon Indian Residential School in southwestern Manitoba was incomplete. She combed through archived death records and identified the names of 70 additional children who had disappeared during the school's active years. Then, in collaboration with the Sioux Valley Dakota Nation—a designated group of Indigenous peoples—she conducted a GPR site survey. The results suggested there were 104 potential graves in three locations.

Later, in 2018, the Muskowekwan First Nation—another designated group in Saskatchewan—asked Terence Clark, an archaeologist with the University of Saskatchewan, for help locating the remains of 35 children who were unaccounted for in school records. A team led by Clark, and including Supernant and Wadsworth, conducted surveys over three large swaths identified by Muskowekwan elders as potential burial sites. The researchers confirmed 10–15 burial sites, which the Muskowekwan First Nation now intends to mark and commemorate.

The Canadian government's Truth and Reconciliation Commission, established in 2008 to document the history and impacts of the residential school system, suspects that there are 139 residential schools across the country with unmarked child burial sites. The potential sites range from coastal wetlands to interior forests and northern permafrost. The subsurface diversity complicates the interpretation of the GPR return signals.

Different terrains have vastly different reflectivity and noisy backgrounds. Picking out a clear, anomalous pattern from a busy background may be challenging, particularly for the small-sized burial of a child. Andrew Martindale, an archaeologist at the University of British Columbia, notes that "anthropologists and archaeologists have been enthusiastic about using GPR, but we've not yet brought sufficient scrutiny to knowing what qualities of the electromagnetic signal correlate with a grave."

Typically, 20th-century cemeteries have a uniform landscape, and graves are dug in loose, homogenous soil with consistent dimensions and spacing. That



THREE-DIMENSIONAL RENDERING of ground-penetrating radar data shows the shape, size, and depth of subsurface buried objects. Some of those objects likely correspond to unmarked burial sites at the Church of the Holy Spirit near Halifax, Nova Scotia, Canada. (Adapted from T. B. Kelly et al., *Forensic Sci. Int.* **325**, 110882, 2021.)

regularity provides a standardized signal for pinpointing unmarked burials. But the early graveyard studies from the 1980s were not generally useful for evaluating precisely how terrain variability affected the return signal that defines a grave. For example, radargrams could not be used to explore how return signals changed with the natural variability from soil types, hydration, and compaction. “There are around 23 variables that people identify as being related to a GPR burial identification, but they depend on the landscape,” says Martindale. He notes that the return signal can vary significantly as a result of those variables and many other near-surface features.

Martindale’s team searches for suspected graves in forested and riverine terrains and other locations unusual for cemeteries. The researchers also study how the return signal changes as a function of season due, for example, to the local water-table depth. One of their findings is that disturbing and then reburying soil—as is done during a burial—results in increased water retention, which in turn produces a subtle signal that can identify a break in the ground’s soil and rock layers and the base of a grave shaft.

Martindale’s basic analysis tools include a set of radargrams for each grave site that consists of virtual slices through the ground, spaced 25 cm apart. His goal is to systematically compile data with that high level of resolution to identify which qualities of the GPR signal are most indicative of a burial.

Refinements in data processing also help to improve the confidence and utility of GPR for identifying graves. For example, petroleum geoscientist Grant Wach and his doctoral student Trevor Kelly of Dalhousie University produce three-dimensional models of underground signal anomalies that indicate unmarked burial sites (see the image above). And on the horizon for GPR is increased use of aerial imagery that can monitor terrains difficult to access on foot. Conducting an aerial search in 2017, UK authorities in northern England found the body of a child murdered in 2004 in the area they suspected the victim to have been buried; the ground was marshy and had glacial deposits overlaid with peat.

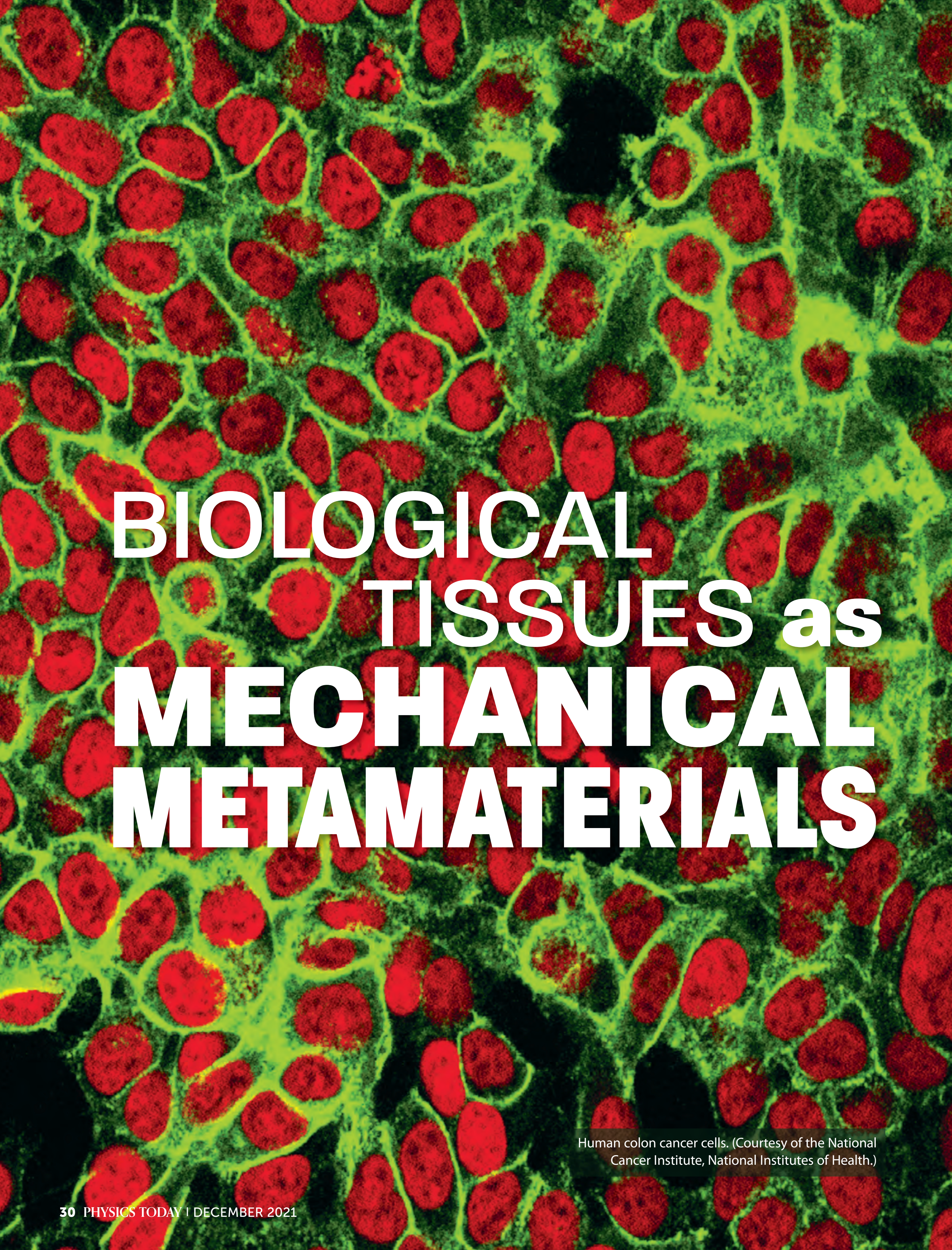
Community involvement

The residential school tragedy was officially remembered in the summer of 2021 with flags lowered to half-mast and memorials consisting of children’s shoes and toys that dotted Canada’s parks and plazas. Orange flags strung across fences and windows in communities across Canada reminded passersby to recall that “every child matters.”

Indigenous communities now want to create grave markers and memorials that tell the bereaved families’ stories. Some are calling for law enforcement to pursue criminal prosecutions. Because many of the communities want to do the investigative work themselves, Wadsworth and his colleagues have begun offering training sessions on the use of GPR systems. Taking pride in the work he and others have done, he notes, “Even grandmothers are coming and wanting to learn about GPR systems. But we know they’re not coming for the technology itself. They’re coming to learn about their past and move to the future.”

Rachel Berkowitz PT

BIOLOGICAL TISSUES as MECHANICAL METAMATERIALS



Human colon cancer cells. (Courtesy of the National Cancer Institute, National Institutes of Health.)

Amanda Parker is a computational scientist at SimBioSys in Chicago, Illinois. She is also the press ambassador for *Biophysical Journal*.

Amanda Parker

How can researchers geometrically tune the extent to which a material embodies a stiff or flexible structure?

When a material is chosen or developed to fulfill a specific function, it transforms from an object into a tool. Concrete, for instance, is designed to support structures, whereas rubber is designed to stretch and bend. Wood is softer than steel but stiffer than nylon. Over time, scientists have amassed, through discovery and invention, naturally occurring and synthetic materials for countless applications. The ones that are chosen depend on the properties needed for the job.

Materials scientists have been so successful in developing and discovering materials that researchers in the field might pause and ask ourselves what our next step is. Should we keep looking for or developing that next material, with exactly the right stiffness for some particular application? Or could we be more clever?

What if we didn't have to pick and choose from our arsenal of materials? What if a material existed that was rigid in certain situations and flexible in others, one that changed exactly when required in order to perform specific functions? Such a material may seem like science fiction, until we realize that many biological systems are, amazingly, able to perform a multitude of tasks by dynamically adjusting their mechanical properties without changing their composition.

During an animal's development, for example, the embryo undergoes morphogenesis, a process in which a collection of interconnected cells—a tissue—reshapes itself (see figure 1). To do that, the tissue must transform from a strong, rigid system to a flowing, fluidlike system in order to take on a new shape, at which point it becomes rigid again.

Unlike normal melting, the transformation can happen at a relatively constant temperature and is well controlled. In the morphogenesis of an embryo, the flowing tissue does not spill and spread randomly but transforms into a particular new geometry, necessary for healthy development. If materials scientists ever hope to be as clever as biology, we will first need to understand what allows such systems to function in that way. But first, we must ask a simple question: What does it mean for one material to be more rigid than another?

Constraint counting

The question of what distinguishes floppy and rigid materials turns out to be one of the oldest questions in physics, and yet it is still not easily answered. James Clerk Maxwell asked the question in the mid 1800s while thinking about macroscopic structures of rods and joints, such as truss bridges.[1]

He observed that a square frame made of four rods would fall over if pushed. If one adds another rod along the frame's diagonal, however, it prevents the frame's collapse, as shown in figure 2a. Adding the rod transforms the frame from something floppy, which moves when pushed, to something stiff that resists motion. One could even say that the frame transforms from liquid-like to solid-like. But how does adding that rod create such a fundamental change in the structure's behavior?

Maxwell explained it by imagining the frame as a collection of degrees of freedom and constraints. The corners act as degrees of freedom because they can move in space, whereas the rods constrain the corners to move only in certain ways. Because any motion of a corner point can be thought of as a set of movements in each spatial dimension—forward and backward, up and down, or left and right—the total number of degrees of freedom is the number of corner points, N_{pts}, multiplied by the number of spatial dimensions, d. Assuming that the four-rod frame can move only in the two-dimensional plane, there exist $N_{pts} \times d = 4 \times 2 = 8$ degrees of freedom. At the same time, there are as many rods as there are points, which yields 4 as the number of constraints, N_c.

Maxwell showed that if one ignores trivial degrees of freedom, such as the translation or rotation of the entire structure, then the difference between the number of degrees of freedom (DOF) and the number of constraints reveals whether the structure is floppy or rigid. In our example, imagine that X is that difference: $X = d \times N_{pts} - N_c$ − (trivial DOF), which becomes $2 \times 4 - 4 - (2$ translational $+ 1$ rotational DOF$) = 1$. Because $X > 0$, the structure is floppy and can collapse. But once a fifth rod is added, the equation changes to $X = 2 \times 4 - 5 - (2$ translational $+ 1$ rotational DOF$) = 0$, and the structure becomes rigid. More precisely, the system is isostatic, meaning that the number of (nontrivial) degrees of freedom exactly equals the number of constraints.

If the value of X contains so much information, it must be important. But what does it represent physically? In addition to the difference between the degrees of freedom and number of constraints, its value represents the number of (again, nontrivial) ways the degrees of freedom (the corners of our frame) can be moved and yet require no mechanical energy. That is, X represents the number of zero modes—ways in which the degrees of freedom can be moved without altering the system's energy. If any such nontrivial zero modes exist, the system is floppy. If none exist, it's rigid.

A little more than 100 years after Maxwell published his work on the stiffness of frames, Christopher Calladine refined the method to take into account "states of self-stress"—special cases in which individual rods are tensed or compressed while the system as a whole is in mechanical equilibrium.[2] Ever since, the resulting simple, powerful Maxwell–Calladine constraint-counting method has been instrumental in mechanical engineering for building sturdy structures that can withstand external forces.

Constraint counting explains why two structures made of the same components, such as a collection of steel rods, can behave quite differently under stress. But how do we explain why steel itself is a rigid material? Remarkably, constraint counting is useful in predicting not only how rigid macroscopic structures can be but also whether microscopic structures, such as a configuration of atoms, will produce a rigid material. In fact, it turns out that the classical solid-state theory of stiffness in an atomic crystal is essentially identical to constraint counting.

In the classical picture, a solid is a collection of atoms arranged in a repeating, crystalline pattern. The structure remains cohesive because each atom interacts with its neighbors through repulsive and attractive forces. Although no rods connect atoms to each other, the interactions among the atoms act as constraints that keep them from getting too close to or too far from each other. Moreover, the interactions are usually short range, so atoms only weakly interact with other atoms that aren't nearby. A representation of that solid—using dots to represent atoms and lines to represent constraints—looks quite similar to the macroscopic frame in the left panel of figure 2b.

Coordination number and jamming

One convenient aspect of a large system, such as a collection of atoms in a crystal, is that researchers can use statistics to analyze it. In a cubic solid, for instance, all atoms but those on the edges of the solid have six neighboring atoms; it would resemble a 3D version of figure 2b. The bulk of the system is so large, though, that the average coordination number—the number of neighbors per atom—is still close to six. Instead of counting the constraints, or bonds, individually in the solid, one can calculate their number, at least on average, by multiplying the number of bonds attached to each atom by the total number of atoms and then, to avoid counting bonds twice, dividing by two.

The number of bonds equals $\frac{1}{2}\langle z\rangle$ times the number of atoms N_{atoms} in the system, where $\langle z\rangle$ is the average coordination number. In three dimensions the number of nontrivial degrees of freedom equals $d \times N_{atoms} - 6$. When the expressions are equal, the system changes from floppy to rigid. That condition thus can be used to find the value of $\langle z\rangle$ that sits at the transition point: $\frac{1}{2} \times \langle z\rangle \times N_{atoms} = d \times N_{atoms} - 6$, for which $\langle z\rangle = 2d - 12/N_{atoms}$. Because $N_{atoms} \gg d$ in a large system, one can ignore the last term and get $\langle z\rangle = 2d$.

So if the average number of neighbors is greater than or equal to twice the dimension, the system is rigid; otherwise, it's floppy. That's a powerful way to think about constraint counting because it means that someone can discern whether a lattice is floppy or rigid simply by knowing the number of neighbors a particle has in the lattice.

Not all materials are composed of a periodic lattice, though. Some of the most interesting and important materials, such as plastic and glass, are disordered (see the right panel of figure 2b). In 1985 theorist Michael Thorpe led a partnership that successfully adapted the ideas of Maxwell and Calladine to amorphous solids,[3] which can have regions of high and low coordination numbers.

That coordination-number perspective sheds light on why some systems seem to spontaneously become solid-like, without any change to their constituent particles, their temperature, or the extent to which the system of particles is disordered. Have you ever tried to pour grains of rice out of the corner of

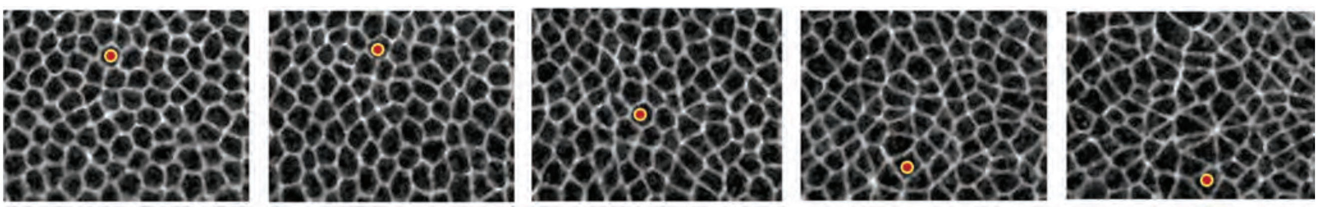


FIGURE 1. EPITHELIAL CELLS in a *Drosophila* (fruit fly) embryo flow during morphogenesis. Each frame in this movie sequence, which proceeds from left to right, is separated by about 1.2 seconds. In the first two frames, the cell marked with a red dot remains roughly stationary in space, as the tissue remains solid-like. Once the tissue begins to flow, the cell moves rapidly downward, as seen in the final three frames. (Adapted from ref. 18; see also H. Evarts, "New View on How Tissues Flow in the Embryo," Columbia University Engineering news release, www.engineering.columbia.edu/press-releases/kasza-tissues-flow-embryo.)

a bag? Or let coffee beans pour from a hopper at the grocery story? In both cases, the particles flow—unless they don't. Sometimes they get stuck, transforming the system from one that behaves like a fluid to one that behaves like a solid. That jamming transition happens when the particle density increases above a critical value[4,5] (see figure 2c).

Mechanical metamaterials

The insights about rigidity suggest an exciting idea: One can control the mechanical properties of a system simply by controlling its geometry. That idea underlies a class of materials called mechanical metamaterials. They are manmade structures whose behavior under external forces depends on the way their components are arranged, not on their composition. Usually that kind of material relies on the geometry of a precisely constructed subunit built into the material and is the counterpart to optical metamaterials (see the article by Martin Wegener and Stefan Linden, PHYSICS TODAY, October 2010, page 32).

By considering only its geometrical properties, researchers can design the material to behave in unusual ways, such as exhibiting auxetic behavior—when the material is compressed in one direction, it also shrinks in the perpendicular direction, as shown in the first two columns of figure 3. It resists fracture and absorbs energy, mechanical properties that are useful in such applications as packing materials, body armor, and shock-absorbing materials.

In designing mechanical metamaterials, researchers have applied Maxwell–Calladine counting in ingenious ways. For example, some have created materials whose zero modes lie on their boundaries, where the average coordination number dips below the critical value, or come into play only when the material is deformed in a particular way. Those materials are actuatable, meaning they change their behavior rapidly upon receiving a specific (in this case, mechanical) signal.[6]

Rigidity in tissues

Materials scientists have gained an understanding about the physical origin of rigidity in everything from sand piles to bridges and opened the door to a whole new class of materials. Seemingly fantastical ideas—a robotic hand, for instance, that can easily deform and flow around an object, only to quickly become rigid again to pick up the object—are now possible because of our understanding of how constituents of a material can jam.[7] To revisit the original focus of this article, it appears that we are making progress toward creating materials that act like biological tissues during the process of morphogenesis and dynamically change their mechanical properties.

But there's a surprising catch. Look again at figure 1 and you'll notice something remarkable. Like a system of rigid rods or atoms, the fruit-fly tissue has some number of degrees of freedom (namely, the cell positions) and some number of constraints (namely, the size and shape of each cell and the fact that the tissue must remain continuous and not open holes in its structure) that restrict the movement of those degrees of freedom.

But unlike other systems of rods or atoms, when the cellularized tissue undergoes changes to its rigidity, its number of degrees of freedom and constraints do not change. That is, none of the coordination number, temperature, or degree of disorder has to change significantly for a tissue to go from an arrested state to a flowing state.

In retrospect, the nature of the transformation is not surprising from a biological perspective. It would be inefficient and difficult for a tissue to often need to change its density—for example, via cell death or proliferation—in such a rapid and precise manner (it does happen, though; see PHYSICS TODAY, June 2017, page 19). In fact, other biological systems, such as networks of collagen fibers, behave similarly. They do not change their connectivity but can undergo changes in stiffness. So counting degrees of freedom and constraints in those systems would not do us any good in figuring out whether they are stiff or soft. What, then, is controlling the rigidity in biological tissues and collagen networks?

Biopolymer-network models

We have more collagen in our bodies than any other protein. The collagen proteins come together to form fibers that then connect to each other to form a higher-dimensional mesh. The resulting collagen network is, in vertebrates, the primary component of the extracellular matrix (ECM), a dense composite of molecules that surrounds cells and tissues and gives them structural support.

Much like our schematics of Maxwell's frame and an atomic solid, a schematic of the ECM also looks like a collection of rods connected at joints. The resemblance makes it again possible to count the number of degrees of freedom and constraints in the matrix in order to estimate the coordination number, or average number of connections per joint. Studies of collagen network images reveal that their average coordination number is around 3.4. That's less than the value needed for structural

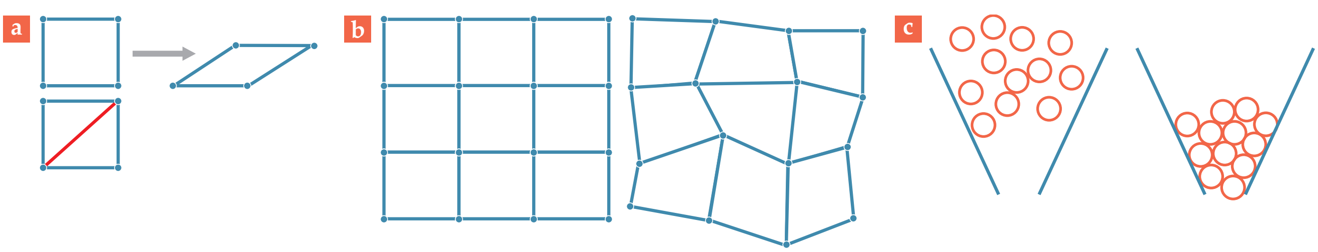


FIGURE 2. FRAMES AND FORMS. (a) With fewer constraints than degrees of freedom, the top frame is floppy and collapses when pushed. With an additional rod (red) on the diagonal, the bottom frame has an additional constraint, such that the structure remains rigid under pressure. **(b)** The crystalline solid on the left has vertices (atoms) embedded in a regular grid. The vertices in the amorphous solid on the right do not lie on a lattice. **(c)** The configuration of particles in the left panel allows them to flow. On the right, the particles are in a jammed, solid-like configuration.

rigidity in both two and three dimensions, as we saw earlier, and it means that those networks should always be expected to be floppy.

Even without changing their connectivity, however—by, say, adding, removing, or rearranging fibers—those networks can transition from weak and floppy to sturdy and rigid. In fact, that ability is biologically important, as many experiments have shown that the stiffness of the ECM acts as a signal to tissue cells. Depending on the ECM stiffness, the cells may be prompted to change or maintain their behavior.[8]

To help understand the mechanical properties of such networks, researchers often use spring-network models, in which collagen fibers are represented as springs connected at points. Indeed, the rod considered in the models above can be thought of as a stiff, unbendable spring. The use of a spring allows us to explore its rodlike limit or to see what happens when connections between points are allowed to stretch or compress, much like real biological fibers. The energy it takes to compress or stretch a spring depends (quadratically) on how far away the spring length is from its preferred, or equilibrium, length.

Using such a model, researchers find that a floppy network becomes rigid when at least some of its springs are not able to reside at their preferred lengths. Importantly, that happens when the network is sufficiently strained. The network does not become more crowded, as in a jammed system, or cooler and more ordered, as in a traditional liquid-to-solid phase transition. Rather, the network experiences a geometric incompatibility. It simply cannot accommodate the newly imposed shape.[9,10]

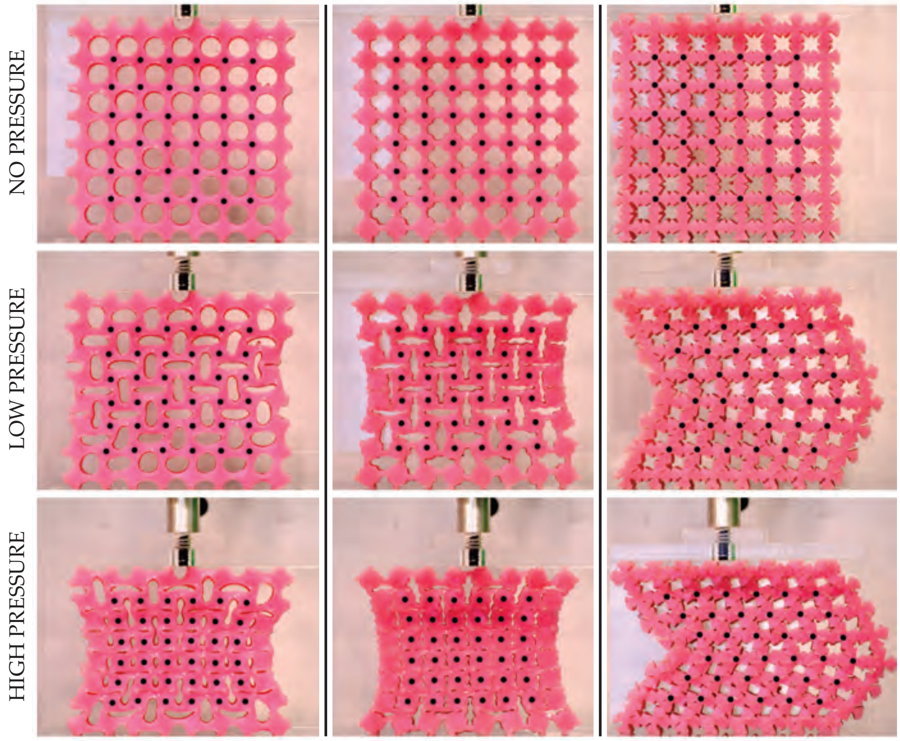


FIGURE 3. MECHANICAL METAMATERIALS. Three mechanical metamaterials made of the same type of rubber but with distinct geometries produce distinct behaviors under compression. Each column represents a structure, initially relaxed (top), with a unique hole shape. Whereas the hole shapes in the first two columns lead to similar behavior under compressive stress, the geometry of the holes in the last column result in a much different final configuration. The center-to-center distance between the holes is 10 mm in the relaxed configurations. (Adapted from J. T. B. Overvelde, S. Shan, K. Bertoldi, *Adv. Mater.* **24**, 2337, 2012.)

Vertex models of tissues

Although intriguing, the emergence of rigidity in collagen networks does not immediately seem applicable to tissues. For one thing, a tissue is a collection of cells, not interconnected fibers. Even so, if physicists are good at anything, it's figuring out how to represent a system as a collection of springlike objects.

One such class of tissue representations is that of vertex models. They describe tissues as—you guessed it—a network of points, or vertices, connected by edges. In this case, though, the polygons created by the vertices and edges represent cells. In many vertex models, it is not the edges that have preferred lengths, as in a spring network, but the polygons (cells) that have preferred shapes.

Real cells can be surprisingly polygonal, with straight, countable sides, and tissues can sometimes even map almost exactly to a special type of vertex model called a Voronoi diagram. The locations of edges and vertices in a Voronoi diagram are determined directly from the locations of the cell centers, as shown in figure 4a. The edges of each cell in a Voronoi diagram enclose the set of points whose distances to that cell's center are less than or equal to the distance to any other.

Furthermore, the idea that cells have preferred shapes comes directly from biology, which tells us that cells, being filled with water and molecules, are fairly incompressible and vary in their elasticity and affinity for sharing edges with other cells. If cells prefer contact with other cells, their edges may be long and their shapes oblong, whereas if they prefer little contact, they are more circular.

In the same way that a floppy spring network becomes rigid when its springs can no longer achieve their preferred lengths, models for confluent tissues—those with no gaps between cells—transform from flowing, fluidlike states to rigid, solidlike states when their cells can no longer achieve their preferred shapes[10–12] (see figure 4c and the article by Ricard Alert and Xavier Trepat, PHYSICS TODAY, June 2021, page 30). Amazingly, researchers have observed that simple, geometric marker of rigidity in experiments on real tissues. In those experiments, the shapes of cells are directly measured and used to correctly predict the tissue's rigidity.[13,14]

Origami and the hunt for global rigidity

Researchers have now established that both collagen networks and confluent tissues have rigidities that can be tuned using

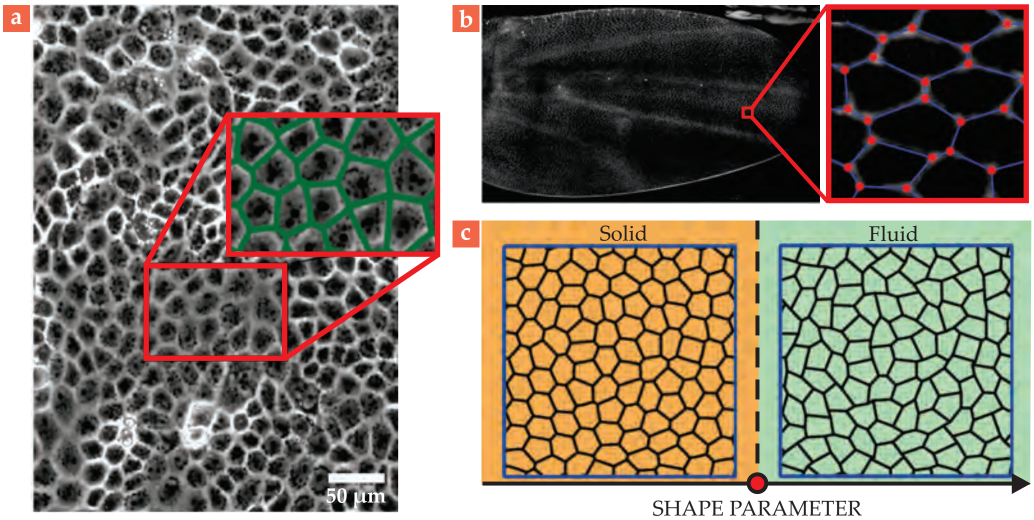


FIGURE 4. VERTEX MODELS for epithelial tissue. **(a)** A layer of tissue showing the cell membranes. The inset pictures a Voronoi tiling (green edges) constructed using the cell centers. The constructed edges, each consisting of points equidistant from the nearest cell centers, almost exactly match the cell boundaries. (Adapted from M. L. Zorn et al., *Biochim. Biophys. Acta Mol. Cell Res.* **1853**, 3143, 2015.) **(b)** A fruit-fly wing during morphogenesis. The inset shows the cell membranes, overlaid with polygonal tiling. (Adapted from M. Merkel et al., *Phys. Rev. E* **95**, 032401, 2017.) **(c)** In this schematic of the vertex-model phase diagram, above a critical value (red dot) for the preferred cell "shape"—the ratio of the average cell perimeter to the square root of the area that the cells would like to achieve—the tissue is fluidlike, and below it is solid-like, despite no change to the tissue's density or coordination number. (Adapted from ref. 12.)

parameters that represent inherent, geometric quantities—fiber length and cell shape, respectively. The process provides a way to design mechanical metamaterials with dynamic mechanical properties. It also brings us one step closer to creating materials that behave like real biological systems.

Yet an overarching question remains: Why doesn't constraint counting work to predict rigidity in those systems? Put another way, can we know when constraint-counting arguments will work and when they won't?

One research field that is providing clues to that mystery is the study of origami. Given a sheet of paper that can be folded only along a predetermined set of lines, how many final, folded configurations exist? Can the paper move freely from one folded state to another, analogous to the way cells in a liquid-like tissue can rearrange and flow? Or is the paper forced to take on one stable configuration, more like cells in a solid-like tissue?

It might seem useful to apply constraint counting to that system because it is composed of edges (folds) and vertices (intersections of folds). But just as in the cases of spring networks and confluent tissues, counting arguments do not correctly predict rigidity. Researchers have now discovered that the zero modes identified via constraint counting are specifically modes that do not affect the constraints to first order in a Taylor series expansion of those constraints. In some cases, however, although the first-order term in the expansion is zero, higher-order terms may not be.

In other words, because constraint counting is capable of predicting rigidity only to first order, it fails in cases where higher-order terms are important. Constraint counting turns out to be a good approximation for rigidity in some cases—which is why it seems to work for them—but in others, it's just not good enough, and one needs to investigate how deformations of the degrees of freedom affect the constraints at higher order.[15]

Mathematicians and physicists are working to figure out exactly when one can use constraint counting and when one needs to go a step further. But already, evidence is showing that the onset of rigidity in confluent tissues may be explained by using higher-order terms in an expansion of the system's constraints.[16]

What's next?

The potential for a new material, whether biological or bioinspired, relies on understanding what truly determines structural integrity across a broad range of systems and in novel environments.

Understanding rigidity has applications in battling diseases. Cancer researchers are learning how important maintaining healthy mechanical properties of cells, tissues, and the ECM is to controlling metastasis.[17] For example, in the image on page 30, thanks to the staining of E-cadherin, a protein on cell–cell boundaries, one can make out cell edges in green. These are human colon cancer cells, and as they become more migratory and invasive, they undergo a transition that is marked, in part, by a change in shape. By beginning to develop a general framework for rigidity, which includes the novel behavior observed in fiber networks and confluent tissues, materials scientists are already producing interdisciplinary discoveries and ideas. Ultimately that will lead us toward healthier, more sustainable lives.

REFERENCES

1. J. C. Maxwell, *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **27**, 294 (1864).
2. C. R. Calladine, *Int. J. Solids Struct.* **14**, 161 (1978).
3. H. He, M. F. Thorpe, *Phys. Rev. Lett.* **54**, 2107 (1985).
4. A. J. Liu, S. R. Nagel, *Nature* **396**, 21 (1998).
5. C. S. O'Hern et al., *Phys. Rev. E* **68**, 011306 (2003).
6. K. Bertoldi et al., *Nat. Rev. Mater.* **2**, 17066 (2017).
7. E. Brown et al., *Proc. Natl. Acad. Sci. USA* **107**, 18809 (2010).
8. J. D. Humphrey, E. R. Dufresne, M. A. Schwartz, *Nat. Rev. Mol. Cell Biol.* **15**, 802 (2014).
9. A. J. Licup et al., *Proc. Natl. Acad. Sci. USA* **112**, 9573 (2015).
10. M. Merkel et al., *Proc. Natl. Acad. Sci. USA* **116**, 6560 (2019).
11. R. Farhadifar et al., *Curr. Biol.* **17**, 2095 (2007).
12. D. Bi et al., *Nat. Phys.* **11**, 1074 (2015).
13. K. E. Cavanaugh et al., *Dev. Cell* **52**, 152 (2020).
14. J.-A. Park et al., *Nat. Mater.* **14**, 1040 (2015).
15. B. G. Chen, C. D. Santangelo, *Phys. Rev. X* **8**, 011034 (2018).
16. L. Yan, D. Bi, *Phys. Rev. X* **9**, 011029 (2019).
17. K. R. Levental et al., *Cell* **139**, 891 (2009).
18. X. Wang et al., *Proc. Natl. Acad. Sci. USA* **117**, 13541 (2020).

PT

Another zooming year

As I write these words, my colleagues and I are preparing articles for this, the final issue of 2021. Last December's issue celebrated the International Year of Sound. This December's focuses on one of the most vibrant areas of modern physics: soft condensed matter.

The *Physics Today* team continues to work from home. I've yet to meet the magazine's two newest editors in person. Despite the challenge of putting together a magazine remotely, we continued to be engaged. Keeping you and your customers abreast of developments in physics and its related sciences continues to be a rewarding job.

Perhaps because of the pandemic, our webinars are becoming increasingly popular. Whenever I moderate one, I scroll down the list of attendees. They come from every continent bar Antarctica. Interest in what physicists are discovering with the tools that you make remains high.

Students have returned to their campuses. Researchers have returned to their labs. For your business before, during, and—we hope will soon be—after COVID-19, *Physics Today* thanks you!

Charles Day,
Editor in Chief

Aerotech, Inc.
American Astronomical Society
American Physical Society
Anderson Materials Evaluation, Inc.
APE Angewandte Physik & Elektronik GmbH
AR Worldwide RF
as mediendesign Inh. Andrea Sindermann e. K.
attocube systems AG
A-VAC Industries, Inc.
AVAK d.o.o.
Axiom Optics
Bartington Instruments, Ltd.
Basic Books
Bluefors Cryogenic Oy, Ltd.
Boston University
Brilliant.org
Bristol Instruments
Circuit Insights LLC
City University of Hong Kong
Cobolt AB
Colorado State University
COMSOL, Inc.
CRAIC Technologies
Cremat
Cryogenic Control Systems, Inc.
Cryomagnetics, Inc.
Duniway Stockroom Corp.
Euclid Techlabs
European Physical Society
Florida State University
Georgia Institute of Technology
ICEoxford, Ltd.
IEEE 2021 Nuclear Science Symp. & Medical Imaging Conf.
Indiana University
INSACO, Inc.
Institute for Advanced Study
Institute of Advanced Science Facilities, Shenzhen
Integrated Magnetics
Intermag Conference
Intermodulation Products AB
Intlvac Thin Film
J.A. Woollam Co., Inc.
Janis Research - A Lake Shore Company
JanisULT
Kavli Foundation
KEYCOM Corp
Keysight Technologies
Kimball Physics
KNF Neuberger, Inc.
Krell Institute
Lake Shore Cryotronics, Inc.
Leybold USA, Inc.
Louisiana State University

to our 2021 Industry Partners

Mad City Labs, Inc.

Massachusetts Institute of Technology

Master Bond

MathWorks, Inc.

MLD Technologies, LLC

Montana Instruments

MPB Communications, Inc.

Nanomagnetics Instruments (Europe)

National Academies of Sciences, Engineering & Medicine

National Institute of Standards and Technology (NIST)

National Reconnaissance Office

Naval Research Laboratory

Nextron Co., Ltd.

Nor-Cal Products

North Carolina Central University

North Star High Voltage

Optica, formerly OSA

OptoSigma Corporation

Oxford Instruments Nanoscience

Pacific Northwest National Laboratory

Park Systems

PDE Solutions, Inc.

Pearl Companies

Pearson Electronics, Inc.

Pennsylvania State University

Pfeiffer Vacuum

Physik Instrumente L.P.

Pittcon

Princeton University Press

Purdue University

Pyramid Technical Consultants, Inc.

Quantum Design (CA)

Quantum Machines

RFHIC Corporation

Rice University

Rohde & Schwarz GmbH & Co. KG

Spellman High Voltage Electronics Corporation

SPIE

Stanford Research Systems

Sutter Instrument

Swabian Instruments GmbH

Synergy Software

Tabor Electronics

Tektronix

The Chinese University of Hong Kong, Shenzhen

The Ho-Am Foundation

The Hong Kong University of Science and Technology

Thyracont Vacuum Instruments GmbH

TOPTICA - USA

ULVAC Technologies

University of Alabama at Birmingham

University of Chicago Professional Education

University of Denver

University of Notre Dame

University of Oregon

University of Southern California

University of Tennessee, Knoxville

University of Wisconsin-Madison

Vescent Photonics, Inc.

Vistec Electron Beam GmbH

Washington University in St. Louis

Wavelength Electronics, Inc.

Western Washington University

Westmont College

Westpress Werbeagentur, GmbH & Co. KG

WITec Wissenschaft Instrumente und Technologie GmbH

Woods Hole Oceanographic Institution

Zhejiang University

Zurich Instruments AG

Confocal microscopy image of banana-shaped colloidal particles with colors added to represent the different orientations of the bananas' short axis. Because the particles vary significantly in curvature and length, they form a globally disordered phase with multiple domains. (Courtesy of Carla Fernández Rico.)

Bas van Ravensteijn is an assistant professor in the department of pharmaceutics, Utrecht Institute for Pharmaceutical Sciences, at Utrecht University in the Netherlands. **Ilja Voets** leads the Laboratory of Self-Organizing Soft Matter in the department of chemical engineering and chemistry and is a member of the Institute for Complex Molecular Systems at Eindhoven University of Technology, also in the Netherlands.

Colloids out of EQUILIBRIUM

Bas G. P. van Ravensteijn and Ilja K. Voets

The fuel-driven assembly of colloids has opened a route to new, biologically inspired active materials.

Spontaneous ordering and formation of superstructures in nature have long inspired physicists, chemists, and materials scientists. The intricate ordering of atomic and molecular building blocks allows for the realization of a seemingly infinite number of different materials, each with its own properties and functions. From hard rocks and gemstones to squishy cells and tissues, all result from a hierarchical organization of their elementary constituents.

That library of structures can be expanded even further by considering not only atoms and molecules but also colloids as building blocks. Colloids—or more correctly speaking colloidal particles—are relatively small entities, with characteristic dimensions of 10–1000 nm, that are dispersed in a continuous medium.[1] Examples of colloidal systems from our daily lives include milk (liquid fat droplets in water), paint (solid particles in water), and smoke (solid particles in air). Despite being significantly larger than atoms and typical molecules, colloidal particles are also subject to Brownian motion that originates from thermal fluctuations of the surrounding medium.[2] They can be considered, therefore, as atoms' and molecules' larger analogues.

Facilitated by their continuous and autonomous movements, colloids can self-assemble like atoms and molecules into larger superstructures. The presence of mesoscopic length scales endows such supracolloidal materials with unique and tunable optical and mechanical properties. For example, adjusting the distribution of colloidal filler particles is a powerful way to increase the strength of otherwise weak organic materials such as plastics.[3] Natural supracolloidal materials include opals and the wings of certain butterflies, in which the periodic arrangement of submicron features generates bright colors (see figure 1; also see "Biomimetics: Lessons on optics from nature's school" by Ross McPhedran and Andrew Parker, PHYSICS TODAY, June 2015, page 32).

Inspired by the analogy between colloids and atoms or molecules and by the prospect of fabricating materials with new properties, condensed-matter physicists and chemists have started focusing on the controlled self-assembly of colloidal particles. Procedures for synthesizing highly uniform and well-defined building blocks have opened a path to myriad new superstructures. They range from close-packed crystals to low-density ordered arrays and finite-sized clusters[4] (see figure 1).

From opals to cells

Although the progress made in the field of colloidal self-assembly is impressive, the functionalities of the resulting materials remain rather limited. That state of affairs becomes more apparent when compared with an average living cell that is able to dynamically reconfigure, adapt, amplify signals, self-replicate, and self-heal (see figure 2). What underlies the mismatch in attainable complexity and functionality? In the case of coarse-grained soft condensed matter, the answer lies in the distinction between thermodynamics and kinetics.

That systems of self-assembling colloids are typically designed based on thermodynamic principles implies that the targeted superstructure exists in or near the global minimum of the free-energy landscape (see figure 3a). In its initial, disassembled state, the system occupies a high-energy state. As the individual particles assemble, the free energy of the system decreases. According to thermodynamics, the structure-forming process is spontaneous.

Because the final structures exist in a deep free-energy minimum, they are stable. The stability is manifested by the absence of major structural rearrangements in response to small external fluctuations. Hence, from the perspective of structural dynamics, typical colloidal self-assembled materials resemble rocks and gemstones rather than self-regulating and adaptive cells.

Luckily, there is more to life than thermodynamic equilibrium. In fact, all living beings rely on self-assembled structures that form or operate far from their thermodynamic ground state. To function, such systems rely on the continuous net exchange of energy or matter with the environment. They are therefore termed dissipative.[5] The dependence on an influx of energy endows the materials with a direct way of interacting with their environment and leads to the fascinating behavior that underlies the transport, motility, and proliferation of cells.[6]

Inspired by that functional richness, the colloidal science community is venturing into the relatively unexplored territories of out-of-equilibrium systems. To imprint active behavior onto colloids, one can follow multiple strategies. For example, external magnetic or electric fields can be used to induce a polarization or magnetization of the particles that triggers their assembly. As soon as the field is removed, the driving force for structuration vanishes and the assemblies disintegrate. Fascinating dynamic structures have been obtained through the temporal control over the strength and direction of the applied fields and over the field susceptibility of the particles.[7]

Alternatively, one can rely on particles that actively propel themselves (see “Microswimmers with no moving parts” by Jeffrey Moran and Jonathan Posner, Physics Today, May 2019, page 44). Those so-called active swimmers are the colloidal analog of rockets that thrust themselves forward in a certain direction. Active swimmers display strongly directional motions that dominate the random, Brownian movements in equilibrium and are typically determined by chemical or thermal gradients in the surrounding medium. Analogous to the flocking of birds and fish, active swimmers can assemble into dynamic superstructures.

In this article we focus instead on a new and different class of out-of-equilibrium colloidal systems. Akin to examples from the biological world, such systems rely on molecular fuels or light to drive their assembly.[8]

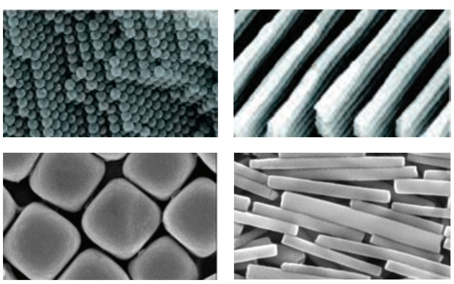


FIGURE 1. MICROSTRUCTURAL ELEMENTS of colloidal length scales give opals (top left) and certain butterflies (top right) their vivid colors. (Adapted from H. Inan et al., *Chem. Soc. Rev.* **46**, 366, 2017.) Synthetic colloidal building blocks of various shapes can also be prepared and assembled. (Adapted from L. Rossi et al., *Soft Matter* **7**, 4139, 2011 [left], and C. Fernández Rico et al., *Adv. Mater.* **31**, 1807514, 2019 [right].)

Microtubules

To illustrate the hallmarks of fuel-driven systems, we take microtubules as an archetypal example.[9] Microtubules are a key structural element of the cytoskeleton, and they are able to reshape and remodel themselves in response to triggers from the outside world. They make highly complex cellular behavior possible.

The tubular structures are formed via the assembly of tubulin dimers. Each tubulin dimer comprises an α tubulin segment and a β tubulin segment, both of which can bind to guanosine triphosphate (GTP). The molecule is an energy-rich fuel, which promotes the tubulin dimers into a self-assembling state (see figure 2). Tubular structures are formed when the α segment of a free dimer binds to the β segment of an adjacent dimer already in the structure. But the chemical stability of the GTP–β-tubulin complex is modest when incorporated in the microtubular structure. That’s because while bound, GTP converts to a lower-energy waste product, guanosine diphosphate (GDP). The resulting GDP–β-tubulin would destabilize the tubular assembly. But so long as the tubular end caps are decorated with the higher energy GTP–β-tubulin complexes, structural integrity can be maintained. Conversely, if the tubulin complexes in the active end cap cannot be kept in the activated GTP form, the end cap will act as an initiation site for the spontaneous and catastrophic disintegration of the microtubules.

In essence, the disassembly of microtubules is driven by the conversion of GTP (fuel) to GDP (waste). Only if the addition of activated tubulin outcompetes the fuel-to-waste conversion in the end caps can fibers be formed. Because the structures’ formation is governed by the kinetics of both tubulin attachment and fuel-to-waste conversion, microtubules are extremely dynamic structures that are able to respond to small fluctuations in the local fuel concentration.

The complete biomolecular machinery involved in regulating the behavior of microtubules is even more delicate and complex than outlined here. Nevertheless, one can discern the basic requirements to imprint its dynamic features on synthetic colloidal materials. Of particular importance is the need to establish a direct link between energy-consuming networks and the assembly process.[5]

The link is depicted in figure 3b. In the equilibrium ground state, the building blocks are in a precursor state; they do not feel a driving force that could arrange them into higher-order structures. The supply of fuel, however, promotes the building blocks to an activated state that makes them mutually attractive. While activated, the building blocks can assemble.

The crucial contrast with the self-assembled structures in equilibrium is that assemblies formed after activation occupy energy states far from the global free-energy minimum. When the fuel runs out, the rate of particle deactivation outcompetes the rate of activation and inactive, non-assembling particles accumulate. Eventually, the assembly disintegrates. The fuel dependence ensures that the assembly behavior of those out-of-equilibrium systems is governed by reaction kinetics rather than thermodynamics, which is the case for equilibrium processes.

Interestingly, the fuel-driven assembly cycle depicted in figure 3b is quite general. Although the deactivation reaction's kinetics need to be tuned to allow for an accumulation of active building blocks and those reactions do not spontaneously occur in the backward direction, there are no mechanistic restrictions. Formation and breakage of covalent bonds—but also of noncovalent interactions or conformational switches—can be used to drive the assembly. That design freedom opens extensive opportunities to design and fabricate synthetic fuel-driven materials. Already, researchers have created an impressive set of out-of-equilibrium materials, such as gels, micelles, emulsion droplets, coacervates, and cell-like vesicles.[8,10]

Despite their generality, the rules of out-of-equilibrium design are tricky to transfer to the colloidal domain. Conceptually, nothing changes. Handling building blocks with colloidal dimensions, however, requires deft control over the interactions between the particles. The interactions can be embodied in the interparticle potential U, which is the sum of repulsive and attractive contributions. Van der Waals forces, which are almost always present, contribute to the attractive part of U.[1] They arise from a mismatch in polarizability of the particles and their surrounding medium.

Because their magnitude scales with particle size, van der Waals forces are much more pronounced for colloidal particles than they are for their molecular counterparts. To prevent strong clustering of colloids, the attractions need to be overcompensated with repulsions. In traditional colloidal systems, the goal is achieved by decorating the surface of the particles with charged moieties or polymer brushes.[1] The resulting electrostatic or steric stabilization ensures the particles remain as well-dispersed single entities in their surrounding medium.

To imprint fuel-responsive behavior on a colloidal system, the fuel needs to temporally switch U from net repulsive to slightly attractive. "Slightly" is crucial here. When the fuel generates attractions that are too strong, the colloidal bonds formed in the activated state are likely irreversible—again because of van der Waals forces, which are extremely strong at short interparticle distances. Once trapped in those deep potential minima, colloids will never—or at least on reasonable time scales—redisperse. Any dynamic, fuel-dependent clustering characteristics are lost.

An additional difficulty associated with the colloidal size of the building blocks is the mismatch in dynamics. Colloidal objects are significantly more sluggish than molecules. When using molecular fuels to guide the assembly of those slower particles, the deactivation reaction's kinetics have to be tuned to keep the particles in the activated state long enough for them to meet each other via slow, diffusive processes. If deactivation precedes encounters, assembly is forestalled, and all supplied energy will be wasted. Bridging those length and time scales becomes increasingly challenging for ever-larger particles. For that reason, most systems reported to date involve relatively small nanoparticles.[11]

Because of those additional challenges, the first generation of out-of-equilibrium colloidal materials started to appear only in the last few years. It goes beyond the scope of this article to discuss all system designs out there. If you want to learn more, please refer to our 2020 review.[11] For a flavor of the state of the art, we highlight two archetypal classes of systems. Each is characterized by the nature of their fuel source.

Small molecules as fuels

Directly inspired by microtubules and other biological systems,

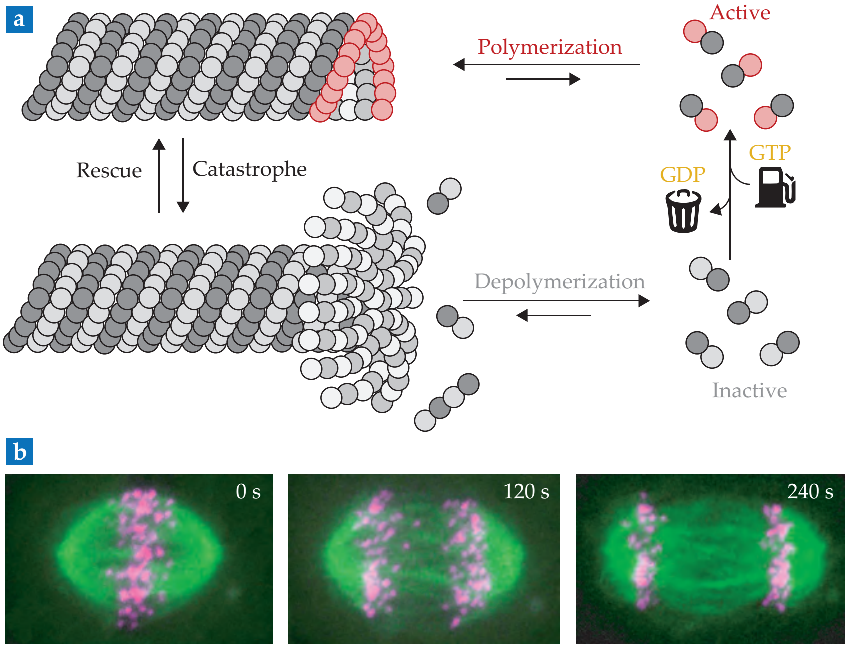


FIGURE 2. POLYMERIZATION OF TUBULIN to form microtubules is fueled by guanosine triphosphate (GTP) **(a)**. Only when one of the two forms of tubulin is in the fuel-activated state (red circles) can polymerization of the elementary dimers into fibers occur. After incorporation, the active building blocks decay to their inactive state, in which GTP converts to the waste product guanosine diphosphate (GDP). The conversion destabilizes the assembly unless the growing end is capped with active dimers. (Adapted from Nordicbiosite.com.) **(b)** Shown here are high-resolution fluorescence microscopy images of microtubules in action during cell mitosis. (Adapted from K. Vukušić et al., *Dev. Cell* **43**, 11, 2017.)

some synthetic colloidal systems run on molecular fuels. An example reported by one of us (van Ravensteijn), Wouter Hendriksen, and our collaborators illustrates the use of molecular fuels to transiently affect colloidal interactions in a striking way (see figure 4a).[12] Our design uses colloids grafted with polymer brushes. In equilibrium—that is, at high pH—the outer brush segments carry negatively charged groups (blue in the figure). The charges generate electrostatic repulsions between the particles, safeguarding their stability.

But when fuel is added, the negatively charged groups are converted to neutral, hydrophobic moieties (red in figure 4a). Because the particles are dispersed in an aqueous medium, their fuel-induced hydrophobic character renders them unstable and causes them to cluster. The trick behind the system is that the hydrophobic groups are metastable only under the imposed conditions. In time, the hydrophobic groups revert to their charged analogues. Once enough charges have accumulated, the cluster spontaneously disintegrates under the action of electrostatic repulsion. Injecting new fuel restarts the assembly cycle.

The cycle cannot be repeated indefinitely because of the accumulation of waste products, which are poisonous to the system. If they remain, they cause irreversible damage to the reaction network by dissolving particles, cleaving nonfuel related bonds, and so on. As soon as the waste is removed, however, new fuel-driven cycles can resume. Waste removal can be achieved through continuous flow, through membrane reactors, or with fuels that generate waste products that spontaneously withdraw themselves from the system through evaporation or precipitation.[13] Biology solved the problem of waste accumulation with elegant regulating mechanisms that use molecular pumps and the coupling of several reaction cycles in tandem. Analogous strategies in the synthetic world could yield materials with extended lifetimes.

Job Boekhoven of the Technical University of Munich and his coworkers have developed a robust, fuel-driven system. In it, precursor particles carry charged groups on their surface that could be temporarily neutralized by the addition of a chemical fuel.[14] The authors showed that different assembled states could be reached depending on the fuel concentration and on the way fuel is supplied over time. That the properties of the final system depend on the fuel's processing or history is unique to out-of-equilibrium materials.

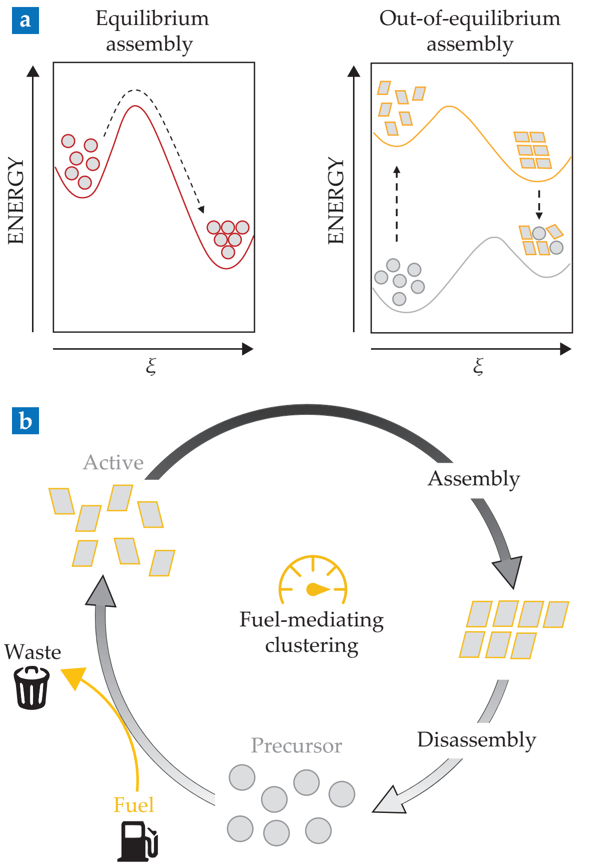


FIGURE 3. SELF-ASSEMBLY can occur **(a)** in equilibrium or out of equilibrium. The two plots represent an energy landscape as a function of the reaction coordinate ξ. In equilibrium self-assembly, the assembled structure occupies a global minimum of the energy landscape and is therefore thermodynamically favored. In dissipative assembly **(b)**, building blocks in the precursor state (gray circles) are promoted into their activated state (yellow rectangles) after consuming fuel. Once activated, their assembly commences. The assembled structure does not exist in a global minimum of the energy landscape and is therefore metastable. Once the fuel is depleted, the activated state of the building blocks can no longer be sustained. The structure spontaneously disintegrates and returns to its equilibrium state of well-dispersed precursors. (Adapted from refs. 10 and 11.)

Light as fuel

Using light to supply energy to dissipative colloidal systems is an alternative strategy. Compared with molecular fuels, using light—or, more precisely, photons—circumvents waste accumulation, thereby boosting the number of assembly cycles that can be performed.

Archetypal light-driven systems rely on photo-induced conformational changes of stabilizing ligands that are immobilized on the surface of the particles.[15] An example of particles functionalized by azobenzene appears in figure 4b. In thermal equilibrium, the azobenzene groups are present in their *trans* conformation (red in the figure). In that conformation, the ligands are hydrophobic and stabilize the particles when dispersed in organic solvents. But when the colloids are illuminated with UV light, the azobenzenes are forced into their metastable *cis* form (blue in the figure). Once in that conformation, a large dipole moment in the azobenzene molecules develops that makes them significantly more polar. Dispersed in what remains an apolar solvent, the dipolar particles are driven to cluster. The energy put into the system by the UV light slowly dissipates as the ligands thermally relax to the *trans* conformation. Eventually, the cluster dissociates.

Relying on light as a fuel also provides a way to regulate where the assembly occurs, because light can be delivered with high spatial resolution. That advantage was neatly illustrated by Bartosz Grzybowski and his colleagues, who fabricated self-erasing images composed of gold nanoparticles.[15] By projecting an image onto a substrate that contained azobenzene-functionalized particles, the researchers induced assembly exclusively in the illuminated areas. Because the nanoparticles are plasmonic in character, their color changes on assembly.

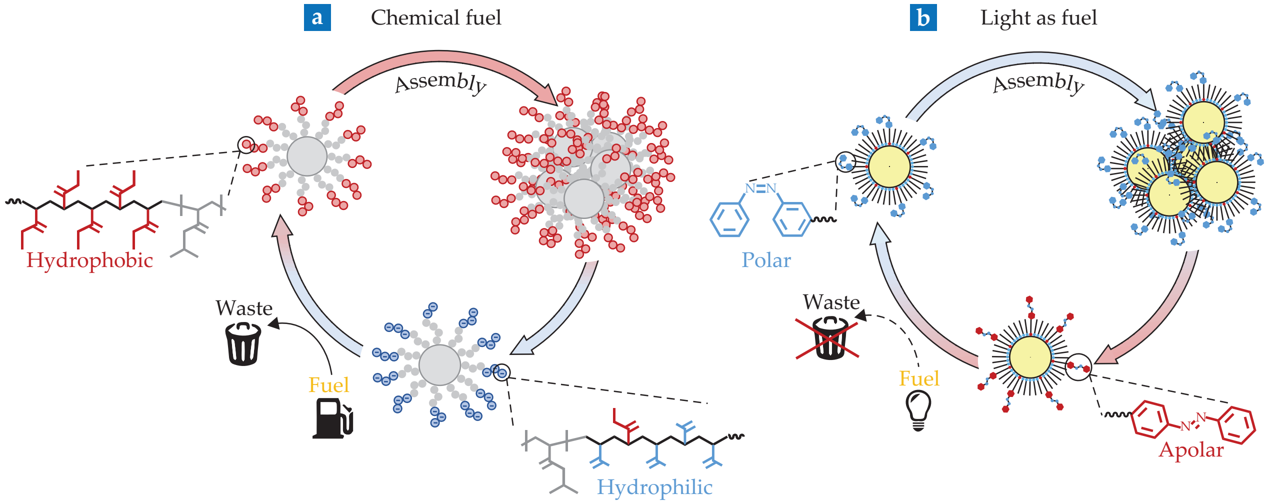


FIGURE 4. COLLOIDAL ASSEMBLY can be sustained by molecular fuel or by light. In the molecular case **(a)**, the equilibrium state (blue) corresponds to the building blocks being charged, stable, and disassembled. Fuel converts the charged moieties to hydrophobic neutral groups (red) and activates them. Once activated, the particles cluster. The hydrophobic groups are not stable in the reaction mixture and slowly convert back to their negatively charged form. Buildup of charges on the particles leads to cluster disintegration. (Adapted from ref. 12.) In the light-driven case **(b)**, thermal equilibrium (red) corresponds to the azobenzene groups that decorate the surface of gold nanoparticles being in an apolar *trans* conformation. On illumination, the azobenzene groups convert to a metastable *cis* conformation (blue). In that conformation, the azobenzene group has a significant dipole moment that renders the molecules polar. Once in the polar form and dispersed in an apolar medium, the particles cluster. In the absence of light, the azobenzenes slowly revert to their *trans* form, thereby removing the driving force for assembly. (Adapted from ref. 13.)

And because the color change is limited to the illuminated areas, an image can be created. Being sustained by light, the assembled state is metastable; the image slowly fades and erases itself.

Into the future

Variations on the two exemplar systems—molecule fueled and light fueled—are being developed at a rapid pace and illustrate the robustness and versatility of the general approach. Of especial interest are systems whose particles are not actively participating in the reaction cycle but respond to a reaction network that runs in the background. Although those reaction cycles are also fueled by small molecules or light-switchable compounds, the tactic eliminates the need to modify the surfaces of colloidal particles through what are often complex synthetic procedures. Reaction cycles that reversibly change pH or surface tension of depletants are just two examples.[11] Additionally, hybrid bio–synthetic systems that combine biomolecular fuel cycles and the tunability of enzymatic or DNA-mediated reaction networks with manmade colloids present a promising future direction.[16]

With those synthetic principles and design rules in hand, it is now up to us and others in the field to start exploring what dissipative colloidal systems can really do. The use of fuels is a promising route to transition from colloidal rocks to colloidal cells. Analogous to the role hard colloidal particles played in elucidating the physics underlying atomic crystallization, the colloidal systems might shed light on the fundamentals of biomolecular out-of-equilibrium processes. What's more, by controlling the interplay between reaction kinetics and particle assembly, researchers can obtain materials that have not just colloidal characteristics but also lifelike ones. Imagine colloidal structures that are self-healing when damaged or that display dynamic oscillating behavior when coupled to periodic oscillations in fuel concentrations.

Toggling between repulsive and attractive states was recently proposed as a route to generate colloidal structures not attainable by straightforward equilibrium self-assembly.[17] Reaching that level of control will entail regulating the relative orientation of the particles in the assembled state and finding fuel reaction networks that perfectly match the colloidal dynamics.

In light of the fast pace by which the field is developing, we have no doubt that those hurdles will be surmounted soon, and we will witness a dynamic revolution in colloidal science.

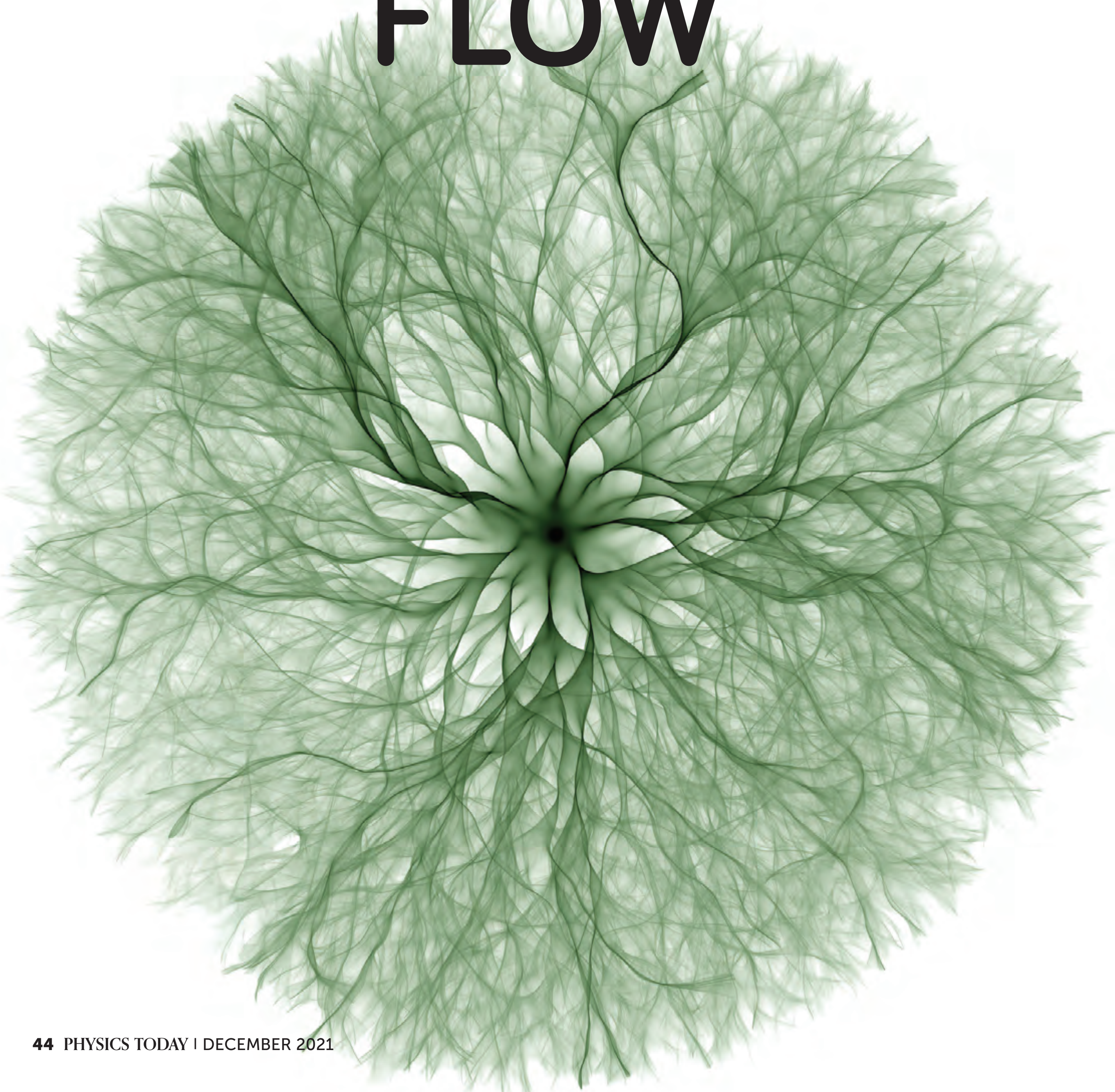
REFERENCES

1. D. H. Everett, *Basic Principles of Colloid Science*, Royal Society of Chemistry (1988).
2. A. Einstein, *Ann. Phys. (Leipzig)* **322**, 549 (1905).
3. L. J. Bonderer, A. R. Studart, L. J. Gauckler, *Science* **319**, 1069 (2008).
4. N. Vogel et al., *Chem. Rev.* **115**, 6265 (2015).
5. G. Ragazzon, L. J. Prins, *Nat. Nanotechnol.* **13**, 882 (2018).
6. E. Karsenti, *Nat. Rev. Mol. Cell Biol.* **9**, 255 (2008).
7. V. Liljeström et al., *Curr. Opin. Colloid Interface Sci.* **40**, 25 (2019).
8. K. Das, L. Gabrielli, L. J. Prins, *Angew. Chem. Int. Ed.* **60**, 20120 (2021).
9. A. Desai, T. J. Mitchison, *Annu. Rev. Cell Dev. Biol.* **13**, 83 (1997).
10. S. A. P. van Rossum et al., *Chem. Soc. Rev.* **46**, 5519 (2017).
11. B. G. P. van Ravensteijn et al., *Langmuir* **36**, 10639 (2020).
12. B. G. P. van Ravensteijn et al., *J. Am. Chem. Soc.* **139**, 9763 (2017).
13. N. Singh et al., *Adv. Mater.* **32**, 1906834 (2020).
14. R. K. Grötsch et al., *J. Am. Chem. Soc.* **141**, 9872 (2019).
15. R. Klajn et al., *Angew. Chem. Int. Ed.* **48**, 7035 (2009).
16. J. Deng, A. Walther, *Adv. Mater.* **32**, 2002629 (2020).
17. M. Tagliazucchi, E. A. Weiss, I. Szleifer, *Proc. Natl. Acad. Sci. USA* **111**, 9751 (2014).

PT

BRANCHED

FLOW



Eric Heller is a professor in the departments of physics and of chemistry and chemical biology at Harvard University in Cambridge, Massachusetts. **Ragnar Fleischmann** is a scientist at the Max Planck Institute for Dynamics and Self-Organization in Göttingen, Germany. **Tobias Kramer** is a professor of theoretical physics at Johannes Kepler University Linz in Austria.



Eric J. Heller, Ragnar Fleischmann, and Tobias Kramer

In many kinds of irregular media, propagating waves enter a beautiful and relatively neglected regime called branched flow. It affects sound, light, water, and matter waves over vastly different length scales.

Small causes can have large downstream effects. That idea is the foundation of chaos theory. But chaos needs time and space to develop, and fascinating behavior can happen on the way. As a directed wave or a collection of rays, such as light, travels through an almost homogeneous medium, even minute random variations in the medium can lead the wave and the rays to branch off and cause extreme intensity fluctuations. That phenomenon is known as branched flow. Examples include the height fluctuations of tsunamis, the flow of electrons in semiconductors, and pulsar radiation propagating through the interstellar medium. This article gives an overview of branched flow, its prerequisites and implications, the breadth of systems in which it has been observed, and the reasons it's so ubiquitous. And the journey to understanding this intriguing phenomenon has just begun.

Branched flow occurs whenever waves travel through complex environments that can be characterized by random, spatially smooth, and modest variations in the refractive index or its analogue. The characteristic spatial scales of the variations (their correlation lengths) must exceed the wavelengths, and their magnitudes must be small enough that the waves or rays representing them are only weakly deflected. In other words, the waves are only forward scattered. Those conditions are frequently met in natural and technological environments. Sound, light, vibrations, and water waves are all dramatically affected by branched flow.

The phenomenon of branched flow shows up in electron waves when they are refracted by residual disorder in high-mobility semiconductors and by deformation potentials in pure metals and semimetals. It emerges in ocean waves deflected by surface eddies in the sea currents; in sound waves refracted in the turbulent atmosphere and refracted underwater by variations in temperature, salinity, and pressure; and in tsunami waves refracted by variations in the ocean's depth. Light undergoes branched flow when it experiences gravitational lensing by galaxy clusters and their associated dark matter, when it is deflected in media with refractive index fluctuations, and when it is refracted by living tissue. Branched flow is even responsible for the voids and filaments in the structure of the universe.

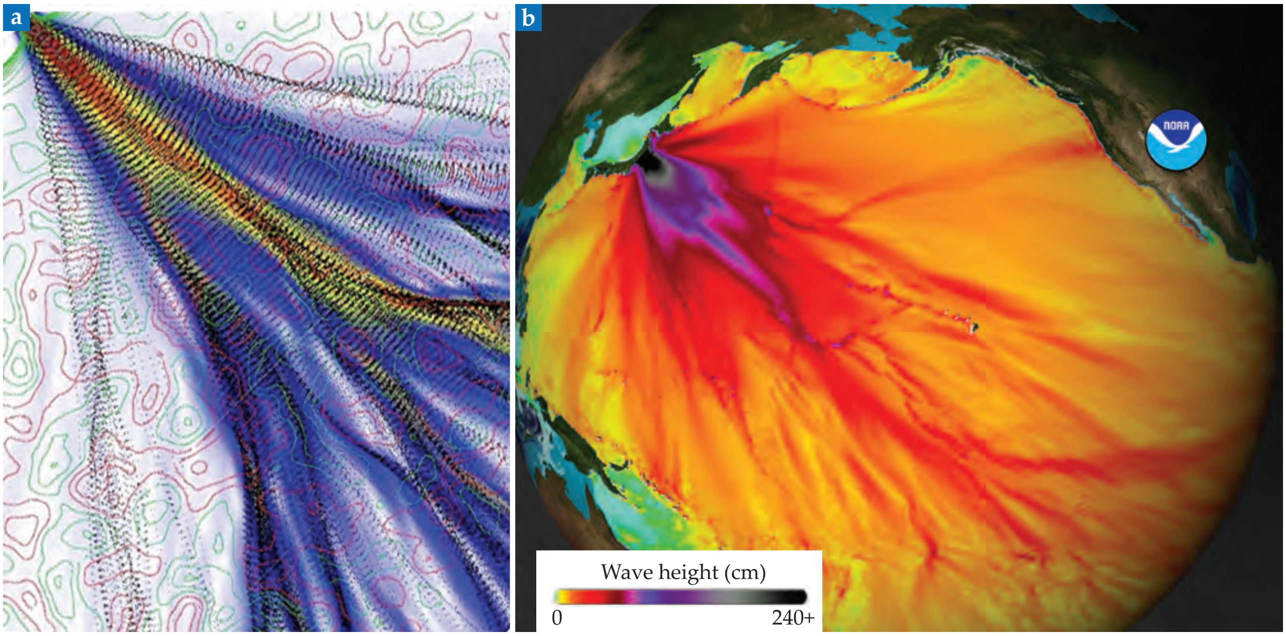


FIGURE 1. BRANCHED FLOW is ubiquitous across different systems and different scales, from semiconductor electron flow to tsunami wave propagation. **(a)** A model shows the micron-scale semiconductor electron flow found by ray (small black dots) and wave (color) propagation from a point source over a random potential field. **(b)** NOAA reconstructed the March 2011 tsunami in the Pacific Ocean. The reconstruction reveals the tsunami's pronounced height fluctuations in the form of a branched pattern. (Image courtesy of NOAA/PMEL/Center for Tsunami Research.)

The behavior has been hiding in plain sight in those and many other physical phenomena. But now awareness of branched flow and its importance is rapidly expanding. The increasing understanding that diverse physical systems with wildly different length scales can reflect similar mathematics and physics is benefiting a wide array of disciplines.

Why is the study of such a ubiquitous and important phenomenon just now getting under way? A likely answer is the lack of computers in the past and physicists' initially slow adoption of computer graphics. Without that computing power, the focus was on analytical theories, which are most feasible at the two extreme ends of the branched-flow problem: the first random focusing events and the final regime of diffusive statistical mixing. Between those extremes lies the ballistic domain of branched flow. The mathematics of that domain remain underdeveloped, and graphical discoveries and computer simulations are essential to present and future development.

Tsunami waves

Perhaps the most terrifying example of branched flow, ocean tsunami waves are a good platform to introduce the phenomenon. Subsurface earthquakes and large coastal landslides can excite highly energetic surface ocean waves with long wavelengths from tens to hundreds of kilometers. The waves travel hundreds of kilometers an hour in shallow water relative to their wavelengths.

The March 2011 magnitude 9 earthquake off the coast of Tohoku, Japan, for example, sent waves propagating out like ripples. For the waves heading away from shore, weak refraction caused by variations in ocean depth accumulated to create dramatic wave height variations in a branched shape, as illustrated in the NOAA tsunami reconstruction in figure 1. That image is typical of branched flow. (To see additional images, including an informative animation, visit https://nctr.pmel.noaa.gov/honshu20110311.)

The speed of the tsunami wave is proportional to the square root of the ocean depth, averaged over scales on the order of the wavelength. Underwater mounds act as focusing lenses,[1] and depressions act as diffusing lenses. Even depth fluctuations of only a few percent can lead to the formation of branches. Unfortunately, the bottom depths of Earth's oceans are not known well enough to accurately predict the locations of distant tsunami branches.[2] If the ocean depths were better

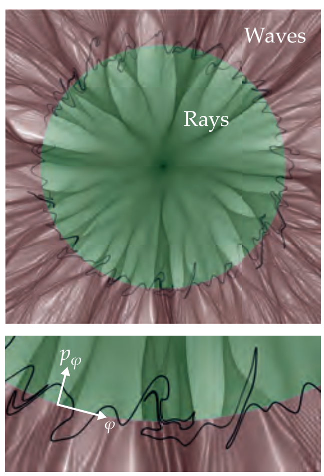


FIGURE 2. CLASSICAL AND QUANTUM branched flow. At top, simulated classical trajectory density (rays, green) and the probability density (waves, purple) flow out from a central point source over the same weak random potential (not shown) at an identical energy. At bottom, a close-up shows the ray momentum tangential to the circle as a function of position on the circle (black line), with the green–purple boundary as an axis. Below the axis, the classical flow is clockwise. In the article's opening image, an extension of the same ray-tracing simulation possesses stable branches (the darkest curving lines). The stability is a matter of chance and will eventually become unstable,[6] as shown in the split of the branch heading upward in the opening image.

known, life-saving predictions could be taken well before destructive energy reached shore.

Because branched flow results from many random small-angle refraction events caused by smooth weak inhomogeneities, such as ocean depth variations, that span more than one wavelength, classical ray tracing applies. For branched flow to occur, the source of the waves needs to be at least weakly localized, such that it can be described as a surface or manifold in phase space. To understand why, imagine how clearly a lens focuses sunlight on a sunny day, when the rays of the Sun are almost parallel, compared with the hard-to-see focus on a cloudy day with diffuse lighting. Examples of localized sources include emission in many directions from a spatially localized point, such as the waves from the Tohoku earthquake and radio waves from a pulsar, and a spatially extended source with restricted initial propagation directions, such as a plane wave (see the box on page 49).

Many natural sources correspond to fuzzy manifolds that are only semilocalized in phase space; examples include light rays from the Sun and waves leaving a storm in one region of the ocean in a range of directions. The branching phenomenon is still rich and dramatic for those averaged sources.

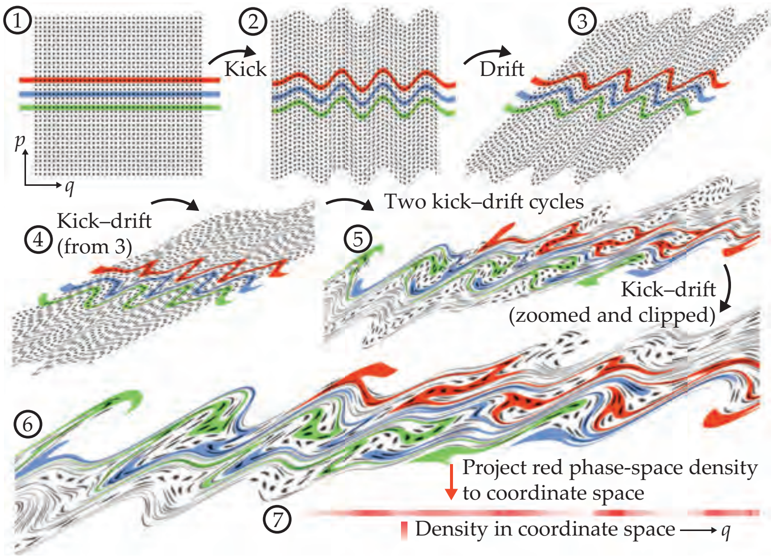


FIGURE 3. A KICK–DRIFT, or thin lens, model implemented in Photoshop shows how branches form. In frame 1, the phase plane starts with circles evenly distributed and uniform colored bands whose widths correspond to ranges of initial momenta. After rounds of momentum kicks followed by periods of drift, many of the original circles get stretched and distorted, whereas others are nearly unchanged, a sign of stable or nearly stable zones. The formation of branches corresponds to the color accumulations in the vertical projection of phase space onto coordinate space, such as shown in the bar in frame 7, which plots the density of the distorted red band.

Ray modeling

The simulations in figure 2 show ray (green) and wave (purple) intensities emitted from a central point source over the same smoothly varying weak random potential. The two solutions have some obvious agreements and some subtle differences. Although the gross branched structures agree, wave interference effects are prominent for coherent sources, such as the frequency-sorted pulsar radio waves arriving at Earth. At long propagation distances, the differences between ray simulations and the actual wave evolution can be profound.

In the opening image, the same ray-tracing simulation is extended to follow more generations of fold catastrophes. It possesses striking stable branches, which result in strong flux density indicated by the darkest curving lines. To track the phase-space characteristics of classical flow, one can draw a surface (in figure 2, the circle) perpendicular to the average flow and plot each ray's position (here the angle φ) of intersection as a function of its momentum component p_φ along that so-called Poincaré surface of section. The intersection (black curves in the figure) reveals how the flow evolves in phase space and indicates when lines fold over, a phenomenon known as caustics. The evolution is equivalent to an area-preserving mapping of the phase plane onto itself.

Photoshop dynamics

A discrete-time nonlinear ray map, known as kick and drift, is a quick and instructive way to reveal branched flow's essence, causes, and effects. The kick–drift ray map replaces propagation across a random potential with a succession of independent and random refractive thin lenses separated by regions in which the wave can flow freely.

To create your own, use Photoshop (or our python script at https://github.com/tobiaskramer/branch) and start with a phase plane (x, p_x) with fuzzy manifolds marked by colored bands that span a small range of initial momenta, as shown in frame 1 of figure 3. Then apply a random area-preserving momentum kick (Filter:Distort:Wave in Photoshop; set the horizontal scale to a minimum), which causes vertical undulations in the manifolds, shown in frame 2. You can adjust the amplitude, number of generators (sines are best), and the wavelength range of the sines by experimentation. (Be careful: You might discover a new regime!) Next, apply an area-preserving drift step (Filter:Distort:Shear), which leads the different momenta from the first step to cause shearing; that step is free-particle drift, shown in frame 3. You can adjust the amount of shearing. Next, repeat that two-step cycle with a random and independent kick–drift, as in frame 4. Just a few iterations lead to frame 5, shown magnified in frame 6.

You can see thickening of the manifold in some regions and thinning in others, even as it is stretched in overall length. Some stable or nearly stable zones survive the six kick–drift cycles, as indicated by the black disks that are still relatively undistorted in frame 6. The points in the red fuzzy manifold end up nonuniformly distributed in position (x) space, as shown in the projection onto coordinate space in frame 7. The red clusters in that projection correspond to branches similar to those in the opening image.

The kick–drift model is equivalent to the small-angle approximation for a series of thin lenses in the optical case. Most zones in frame 7 have contributions from more than one red region in the phase plane. In the wave or quantum version, those separate contributions would carry their own phase and amplitude and interfere constructively or destructively with

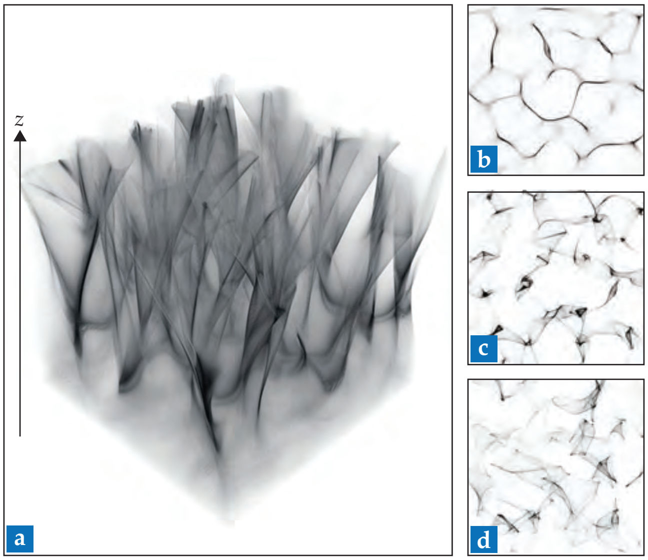


FIGURE 4. THREE-DIMENSIONAL flow from classical ray tracing branches as it propagates through a medium with modest random correlated refractive index changes. **(a)** The flow begins at the base and is uniform in density and with every ray heading vertically. **(b–d)** Two-dimensional slices reveal how the flow accumulates into strong tubes or branches.

one another. The qualitative features here apply even after thousands of kick–drift cycles.

The kick–drift model can be defined as a simple point-to-point area-preserving mapping of the phase plane onto itself, under

$$p_{n+1} = p_n - \left.\frac{dV_n(x)}{dx}\right|_{x=x_n} \text{ (kick step) and}$$

$$q_{n+1} = q_n + p_{n+1} \text{ (drift step).}$$

The potential $V_n(x)$ changes randomly with each iteration n, although it retains certain statistical properties, such as a correlation length in x.

Dimensionality

The branching phenomenon applies in both two- and three-dimensional wave flow. Figure 4 is a ray-tracing simulation of 3D propagation similar to what might occur for sound waves refracting in an atmosphere with temperature and velocity fluctuations.

A common real-world example is the irregular variations in the volume of a jet aircraft overhead. They occur because the loudness pattern on the ground, akin to the light pattern at the bottom of a pool, moves as the jet moves. Indeed, the chance that atmospheric focusing of damaging sonic booms could affect such sensitive places as operating rooms contributed to the 1971 ban of commercial supersonic flight over land in the US.

Characteristic length scales

In figure 2, what determines the average distance between the first focal cusps along the angular direction and the average radial distance L from the origin to the first cusps?[3] For classical flow of energy E traveling over smooth potential undulations of typical height or depth $\epsilon \ll E$, the distance between cusps is $L \propto d(E/\epsilon)^{2/3}$ in terms of the correlation length d of the potential bumps.

That and other statistical properties of the branched flow are largely independent of the details of the random potential and are thus as universal as the concept of a mean free path.[4] Over an ensemble of random potentials, the momentum starts as a delta peak at $\varphi = 0$ and then diffuses as

$$P(\varphi,t) = \frac{1}{\sqrt{4\pi Dt}}\, e^{-\frac{\varphi^2}{4Dt}},$$

with $D \propto \epsilon^2/E^{3/2}d$. The direction of travel decays as

$$\int P(\varphi,t)\cos(\varphi)d\varphi = e^{-t/\tau},$$

where $\tau = 1/D$.

To have branched flow in a system, its dimensions must be smaller than the mean free path, or, at least, the observation time must be shorter than the mean free time. Beyond the mean free path, rays will start to turn around, and their propagation will no longer be ballistic but diffusive. But in weakly refracting media, when $\epsilon/E \ll 1$, the mean free path $l = \langle|v|\rangle\,\tau \propto (\epsilon/E)^{-2}d$ is much larger than the typical branching length scale, which scales like $(\epsilon/E)^{-2/3}$. Thus for a wide regime, wave propagation is dominated by branched flow.

As mentioned earlier, branched flow has two other prerequisites. First, the source needs to be restricted in phase space—for example, a point, parallel ray, or plane wave. Sources with fuzzy but still somewhat localized manifolds, however, also produce pronounced branched flows and lead to extreme events. The second prerequisite is that the flows' wavelength needs to be smaller than the correlation length of the random medium.

Stable branches

More study is needed to elucidate the geometry of classical ray-tracing branched flow, the geometry of wave-propagation branched flow, and the relation between them. For example, caustics—phase-space fold catastrophes projected onto coordinate space—do not alone account for the formation of strong branches. Another contributing factor is quantified by the rarefaction exponent, a measure of the stretching or compression of the initial manifold tangential to the manifold surface.[5] Compression along that direction can pile up flux density and create branches without caustics or enhance branches with them.

In one of the first numerical studies of the branched flow of sound in the ocean, Michael Wolfson and Steven Tomsovic of Washington State University in Pullman recognized a stable subclass of branches: zones of initial conditions in which trajectories coincidentally remain nearby one another for some finite distance or time.[6] For some initial conditions, successive dilations and contractions of phase space almost temporarily cancel each other. The odds of such lucky initial conditions decrease exponentially with time, but some nearly stable zones continue to exist. Examples of nearly stable zones are the relatively undistorted disks in figure 3 that survived six violent kick–drift steps and the ray bundles in the opening image.

Structure of the universe

Caustics and folds contribute to the distribution of matter in the universe. The popular Zeldovich model of the large-scale

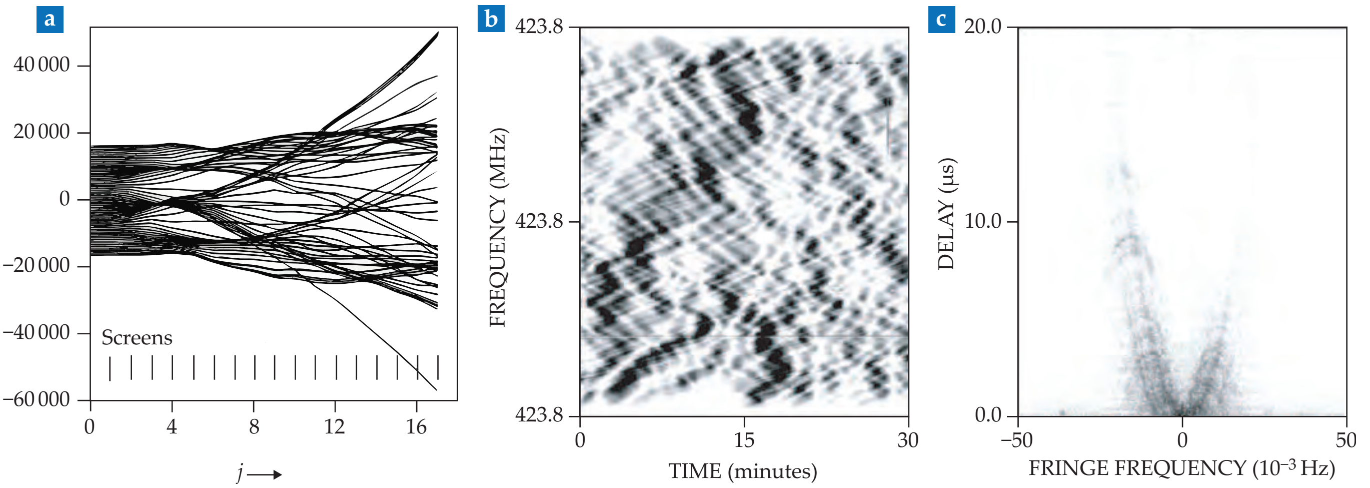


FIGURE 5. OUR GALAXY may cause branched flow. **(a)** This ray tracing represents a model of radio-wave propagation through refracting interstellar clouds. These rays are one of the earliest depictions of branched flow. (Adapted from ref. 9.) **(b)** This graph shows radio-signal strength as a function of frequency and time from pulsar B0834+06, as measured at the Arecibo Observatory in Puerto Rico. The pattern may result from wave interference from branches. **(c)** The double Fourier transform of the pulsar B0834+06 data looks very different from the original spectrum. These sharp parabolic arcs reveal properties of the interstellar clouds. (Panels b and c adapted from ref. 11.)

structure of the universe proposes that the universe resulted from a single kick–drift episode, followed by sticky gravitational effects after caustics formed. According to the model, the relatively uniform distribution of the matter in the early universe possessed random fluctuations in velocities down to some cutoff length scale. (An alternative explanation is that the initial dynamics resulted from a slow gravitational acceleration because of nascent mass-density fluctuations.)

Both mass and velocity variances could arise from quantum fluctuations in the first moments of the universe. Cases with mass, velocity, and combined mass and velocity variations include a long period of free or weakly accelerated drift of relatively tenuous and mostly noninteracting matter. That kick–drift cycle leads to the formation of caustics, which resemble those in figure 4b, before gravitational effects change the dynamics. Numerical simulations based on those ideas give rise to structures similar to the observed cosmic web of matter, including the large voids,[7] as highlighted in the "Caustics and pancakes" section of James Peebles's classic 1980 textbook *The Large-Scale Structure of the Universe*.

The initial kick–drift episode is analogous to sunlight falling on a pool bottom after being refracted just once by surface waves. How matter flies off the undulating surface of a comet is much the same story: Matter is ejected mostly in the normal direction from every point on the comet's surface.[8]

Pulsar radio waves

A similar model applies to radio waves emitted by pulsars. As the waves make their way to Earth, one or more partially ionized

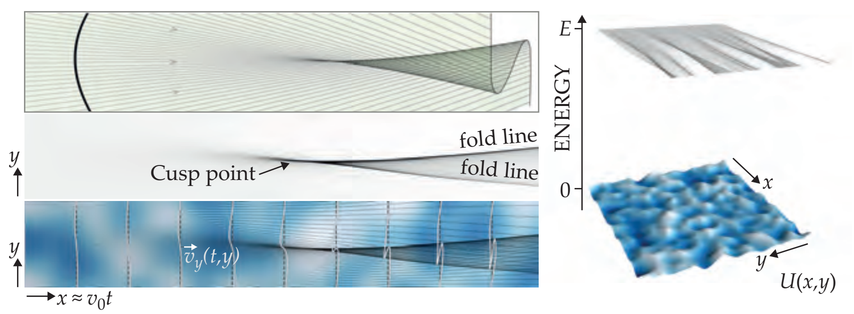
FIRST STEPS IN BRANCHED FLOWS

Branched flow is the spatial pattern that forms when a wave or a bundle of rays is launched over a smooth, weakly refracting random potential. The pattern begins (but doesn't end) with singularities called cusps, shown here. In optics, such singularities are called caustics and typically result from refraction or reflection by curved surfaces.[21] Along caustics, the number of ray solutions passing through a point in coordinate space changes abruptly.

In the top left panel, the rays pass from left to right perpendicular to a concavely curved surface (bold black curve). The result is an imperfect focus: a cusp caustic from which two fold caustics sprout (see the middle left panel). The geometry resembles that of a folded sheet.

The bottom left panel shows how initially parallel trajectories (thin black lines) are weakly deflected by the potential indicated by a blue-scale color map in the background. The white lines show successive phase-space representations (y versus v_y) of the ray bundle. Although for short times only one ray solution passes through every y, after the cusp point a region emerges in which the white line traverses each point three times.

The top and middle left panels show that the projection of the phase-space manifold on to real space leads to singularities in ray density. On the right, an initially parallel bundle of trajectories with total energy E traverses along the x-axis over a correlated random potential $U(x,y)$ with mean $\langle U \rangle = 0$ and standard deviation $\sqrt{\langle U^2 \rangle} = \epsilon = 0.01E$.



interstellar clouds often refract them. Between clouds, the waves propagate in long expanses of free space until finally reaching Earth. The radio source is ideal for observations: pulsed, broadband across frequencies ω, and a virtual pinpoint in the sky. The ionized clouds disperse the waves with propagation speed proportional to $1/\omega^2$ and total time delays that scale with the total density of electrons in the wave's path. For a typical pulsar at 1 kpc distance from Earth, microwaves take several minutes or more to arrive, with the shortest wavelengths arriving at Earth first. The refractive index in a cloud likely varies at many different scales.

Instruments installed at radio telescopes, such as the Arecibo Observatory in Puerto Rico, process data to produce what's known as a spectrogram. Such measurements track the signal strength as a function of time, typically on the order of hours, and frequency, typically a snapshot of a 10 MHz interval between 30 and 1500 MHz, well within uncertainty principle limits.

Besides the characteristic pulsations, whose periods lie mostly in the range 0.1 to 10 seconds, a typical signal varies dramatically by the day, hour, and minute at a single telescope. That variation results from a combination of the pulsar moving relative to a cloud, Earth presumably passing through different branches, and the interference of coherent microwaves that have taken different paths to reach the telescope.

Pulsar radio waves are the earliest example that we (the authors of this article) could find of ray tracing leading to a clear depiction of the branched-flow regime, as shown in figure 5a. Alexander Pidwerbetsky introduced the technique in his 1988 dissertation,[9] which built on his seminal pulsar paper with James Cordes and Richard Lovelace.[10]

A recent innovation uses a double Fourier transform of the radio telescope time–frequency dynamic spectrum, or primary spectrogram, to create a secondary spectrum, or Fourier spectrogram.[11] The results look nothing like the primary spectrum, shown in figure 5b, and contain information about the properties of the interstellar clouds, such as their locations, structures, and densities. They often feature one or two sometimes sharp parabolic arcs, characteristic of each pulsar, as shown in figure 5c.

The clouds are modeled as thin sheets and can generate something akin to kick–drift dynamics for several successive cloud encounters. As early as 1986, Cordes and Aleksander Wolszczan noted, "multipath refractive scattering must be a common occurrence" in pulsar microwaves arriving at Earth.[12] That is, they identified branched flow.

The pulsar data are so rich that they inspired new laboratory measurements on other systems—in particular, experiments with broadband and pulsed point sources and detectors with time scales shorter than those in the medium. Such probes are well suited for measuring turbulence.

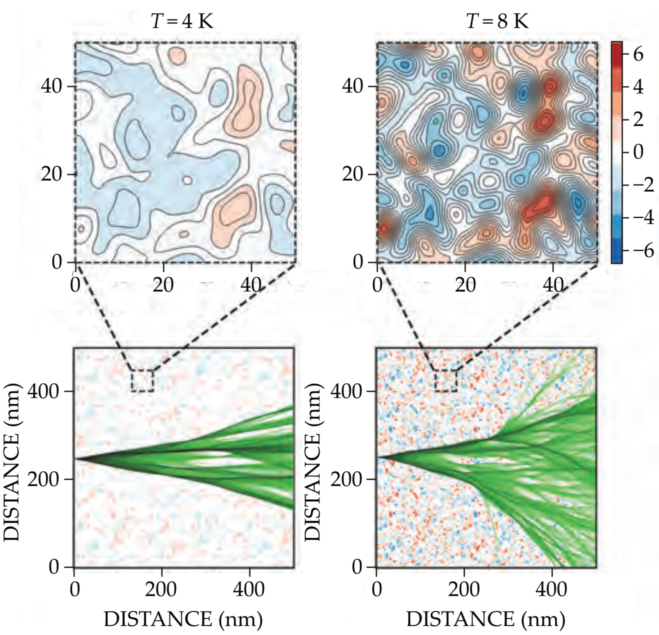


FIGURE 6. CONDUCTION ELECTRONS experience forces as the gradient of the metallic deformation potential, shown at different temperatures in the top two contours. At bottom, semiclassical electron pathways reveal branched flow typical of weak random scattering. Trajectories need to interact with many bumps to produce branched flow. The potential at 8 K has twice the magnitude and half the correlation length as the one at 4 K, so branching happens faster.

Semiconductors and pure metals

We first encountered branched flow in micrometer-scale scanning-probe-microscope measurements of electron flow in 2D electron gases (2DEGs). About 20 years ago, Robert Westervelt's lab at Harvard University produced spectacular images of that electron flow by mapping out the variations in transmission through the 2DEG. Those images include high-resolution ones showing interference fringes in the scattered electron waves.[13,14] (See the article by Mark Topinka, Bob Westervelt, and Rick Heller, PHYSICS TODAY, December 2003, page 47.) Theoretical models of a typical potential field experienced by an electron in a 2DEG revealed the origin of the branched flow in those experiments.[14] Adjacent ionized atoms supply electrons to the 2D layer and induce a smooth nonuniform potential field that randomly deflects the electrons flowing through a small opening between two electrodes, known as a quantum point contact.

Recently, researchers established theoretically that branched flow governs electrical resistivity in pure metals and semimetals from 0 K up to room temperature and often beyond. At finite temperatures, lattice vibrations in a material cause strain and a resulting deformation potential. The researchers treated that deformation potential as exerting a classical force on the conduction electrons. Previously, the deformation potential had been included only in the context of first-order, single-phonon quantum perturbation theory. The electrons could change directions and thus cause scattering and resistivity only by creating or annihilating a quantized lattice vibration, or phonon. We treated successive scatterings with the Boltzmann equation.

The transition from a low-temperature T^4 rise in resistivity in 2D systems and a T^5 in 3D ones to a T^1 rise at high temperatures was seen numerically and analytically using branched-flow trajectories. Classical perturbation theory finds the same expressions, including prefactors, for the resistivity as quantum perturbation theory. The branched-flow description introduces a new picture of electrons colliding with moving potential bumps, which exchange momentum and energy, and the ability to go beyond the perturbative regime. That picture, il-

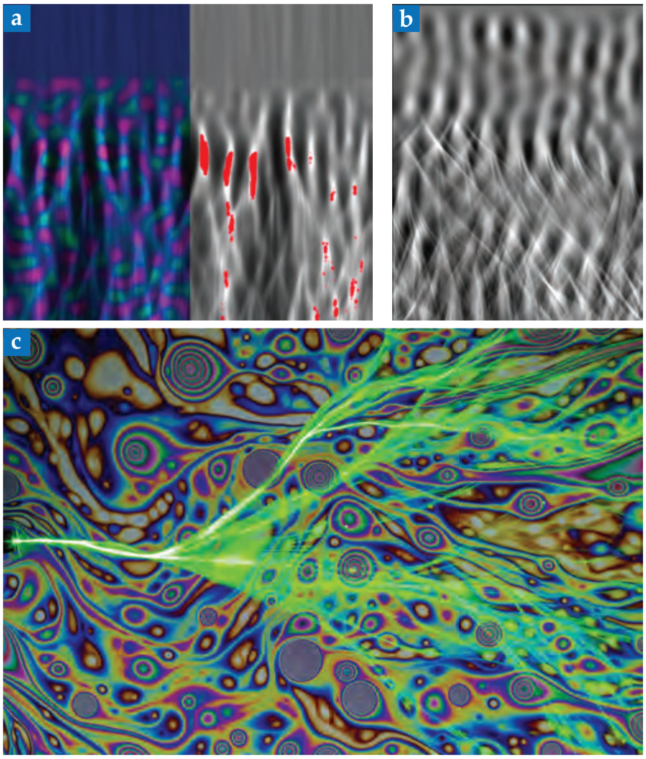


FIGURE 7. FREAK WAVES, fuzzy manifolds, and laser beams. **(a)** In this simulation's time-averaged energy, waves (grayscale, at right) start from the top and propagate in directions that vary by up to 10 degrees and with random phases. When they impinge on a random potential field (left), large-amplitude freak waves emerge (red, at right) with characteristic length scales of more than a wavelength. **(b)** In the evolution of a fuzzy manifold after many kick–drift cycles, branched flow shows not only strong contrast but increasing detail and contrast. **(c)** A laser beam (green) injected into a soap bubble film from the left has branching beams. White-light illumination reveals the background interference pattern caused by variations in the bubble thickness.[20] (Image courtesy of Shruti Saiji, Anatoly Patsyk, Uri Sivan, Mordechai Segev, and Miguel Bandres.)

lustrated in figure 6, captures what is really going on rather than just statistical measures.

Freak waves and soap bubbles

Shortly after the 2DEG electron flow data were understood to result from branched flow, the search began for other wave phenomena that might experience a branched-flow regime. That search revealed that branched flow is a universal wave phenomenon. An important example is the propagation of a storm's wave energy through random refracting current eddies in the ocean. With one exception, the field of extreme ocean waves had made a jump from a theory of uniform sampling Gaussian random statistics, developed by Michael Longuet-Higgins in the 1950s, to considerations of nonlinear wave interactions. The impetus for that jump was the field observations, starting in the 1990s, of far more freak wave events than accounted for in uniform Gaussian random statistics.

Benjamin White of Exxon Research in New Jersey and Bengt Fornberg of the University of Colorado Boulder were the exception; they attributed freak waves to focusing by gyres. Their 1998 study using a simple incident plane wave gave a clear early example of caustics and branched flow.[15] But they did not consider dispersion in the initial wave propagation direction—in other words, they used a sharp initial manifold rather than a more realistic fuzzy one. Critics in the field argued that fuzziness would wash out the effects of caustics. Subsequent work showed that initial dispersion does not wipe out even large wave energy fluctuations downstream. In fact, the branching develops more contrast and even finer structure downstream (see figures 7a and 7b).

Schemes that include those fluctuations in nonuniform Gaussian sampling predict a factor of 50 more freak wave events than those with uniform sampling; no nonlinear wave interactions are required.[16] (Nonlinear effects are, however, important in the formation of the largest waves.[17])

The connection between branched flow and freak ocean waves has inspired studies of freak-wave formation in microwave cavities[18] and of methods for selectively populating optical branches.[19] Visible light injected into a soap bubble film can also undergo branched flow if the variations in the bubble-film thickness are random and correlated. Figure 7c shows an image from one such experiment.[20]

Branched flow emerges everywhere. It lies in the interesting regime between the first focal cusps and eventual random diffusive scattering. The universality of the phenomenon and the huge range of applications at astronomically different scales, from nanometers to the whole universe, suggest branched flow is worthy of further investigation with numerical, experimental, and, especially, mathematical characterizations.

We thank Lev Kaplan, Hans-Jürgen Stöckmann, Scott Shaw, Jiri Vanicek, Jakob Metzger, Henri Degueldre, Erik Schultheis, Gerrit Green, Anna Klales, Byron Drury, Robert Lin, and Steve Tomsovic for valuable collaborations and discussions.

REFERENCES

1. M. V. Berry, *Proc. R. Soc. A* **463**, 3055 (2007).
2. H. Degueldre et al., *Nat. Phys.* **12**, 259 (2016).
3. V. A. Kulkarny, B. S. White, *Phys. Fluids* **25**, 1770 (1982); L. Kaplan, *Phys. Rev. Lett.* **89**, 184103 (2002).
4. J. J. Metzger, R. Fleischmann, T. Geisel, *Phys. Rev. Lett.* **105**, 020601 (2010); *Phys. Rev. Lett*. **112**, 203903 (2014).
5. E. J. Heller, S. Shaw, *Int. J. Mod. Phys. B* **17**, 3977 (2003).
6. M. A. Wolfson, S. Tomsovic, *J. Acoust. Soc. Am.* **109**, 2693 (2001).
7. S. N. Gurbatov, A. I. Saichev, S. F. Shandarin, *Phys.-Usp.* **55**, 223 (2012).
8. T. Kramer et al., *Adv. Phys.: X* **3**, 1404436 (2018).
9. A. Pidwerbetsky, "Simulation and analysis of wave propagation through random media," PhD thesis, Cornell U. (1988).
10. J. M. Cordes, A. Pidwerbetsky, R. V. E. Lovelace, *Astrophys. J.* **310**, 737 (1986).
11. D. R. Stinebring, *Chin. J. Astron. Astrophys.* **6**, 204 (2006).
12. J. M. Cordes, A. Wolszczan, *Astrophys. J.* **307**, L27 (1986), p. L30.
13. K. E. Aidala et al., *Nat. Phys.* **3**, 464 (2007).
14. M. A. Topinka et al., *Nature* **410**, 183 (2001).
15. B. S. White, B. Fornberg, *J. Fluid Mech.* **355**, 113 (1998).
16. E. J. Heller, L. Kaplan, A. Dahlen, *J. Geophys. Res. Oceans* **113**, C09023 (2008).
17. L. H. Ying et al., *Nonlinearity* **24**, R67 (2011); G. Green, R. Fleischmann, *New J. Phys.* **21**, 083020 (2019).
18. R. Höhmann et al., *Phys. Rev. Lett.* **104**, 093901 (2010).; S. Barkhofen et al., *Phys. Rev. Lett.* **111**, 183902 (2013).
19. A. Brandstötter et al., *Proc. Natl. Acad. Sci. USA* **116**, 13260 (2019).
20. A. Patsyk et al., *Nature* **583**, 60 (2020).
21. M. V. Berry, C. Upstill, in *Progress in Optics*, vol. 18, E. Wolf, ed., North Holland (1980), p. 257; J. F. Nye, *Natural Focusing and Fine Structure of Light: Caustics and Wave Dislocations*, IOP (1999). PT

CONESLAYER/WIKIMEDIA COMMONS/CC BY 3.0

Palomar Observatory, home to the 200-inch Hale Telescope pictured here, is one of the most important astronomical research sites in history.

A tale of two telescopes

Home to the 200-inch Hale Telescope and the 48-inch Samuel Oschin Telescope, Palomar Observatory in Southern California has been at the forefront of astronomical research for more than seven decades. The Hale Telescope was the largest such instrument in the world when it saw first light in 1949, and although in recent years larger telescopes have been built at sites with better atmospheric conditions, it has probably contributed more to our understanding of the universe than any telescope other than Galileo Galilei's.

The Oschin Telescope, which began operation in 1948, has a similarly impressive resumé. The wide-field survey telescope's first decade in operation was devoted to the National Geographic Society–Palomar Observatory Sky Survey, a survey of the northern sky that led to the discovery of innumerable astronomical objects and was supplanted by digital successors only in the last 15 years.

Nevertheless, Linda Schweizer's new book, *Cosmic Odyssey: How Intrepid Astronomers at Palomar Observatory Changed Our View of the Universe*, is not a history of the telescopes themselves, nor of Palomar Observatory (although there's plenty of history in it). Her goal is more ambitious: to describe a selection of the century's most important advances in astronomy from the perspective of Palomar Observatory.

That approach is uniquely suited to Palomar. The observatory has been responsible for the discovery of an astonishing variety of astronomical objects, including the first quasar, the first brown dwarf, the first Centaur asteroid, the first solar-system object with an orbit extending beyond the Kuiper belt, and the supermassive black hole recently imaged by the Event Horizon Telescope collaboration.

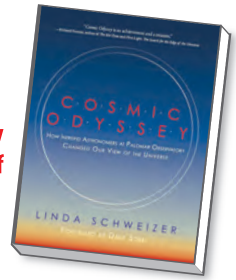
Less spectacular but equally important results have come from long programs of observations at Palomar. Allan Sandage described the evolution of stars after they exhaust the hydrogen fuel at their centers in 1953 and worked throughout his career to refine the cosmic distance ladder. Fritz Zwicky hypothesized that the tidal forces at play during encounters between galaxies create galactic filaments and other peculiar structures. Other developments include the interpretation and classification of supernovae, which led to the discovery that the expansion of the universe is accelerating; the first description of the evolution of quasars over cosmic time; the study of young galaxies at high redshifts and of the intergalactic gas from which they formed; and the first influential models of the Milky Way's formation.

Cosmic Odyssey

How Intrepid Astronomers at Palomar Observatory Changed Our View of the Universe

Linda Schweizer

MIT Press, 2020. $39.95

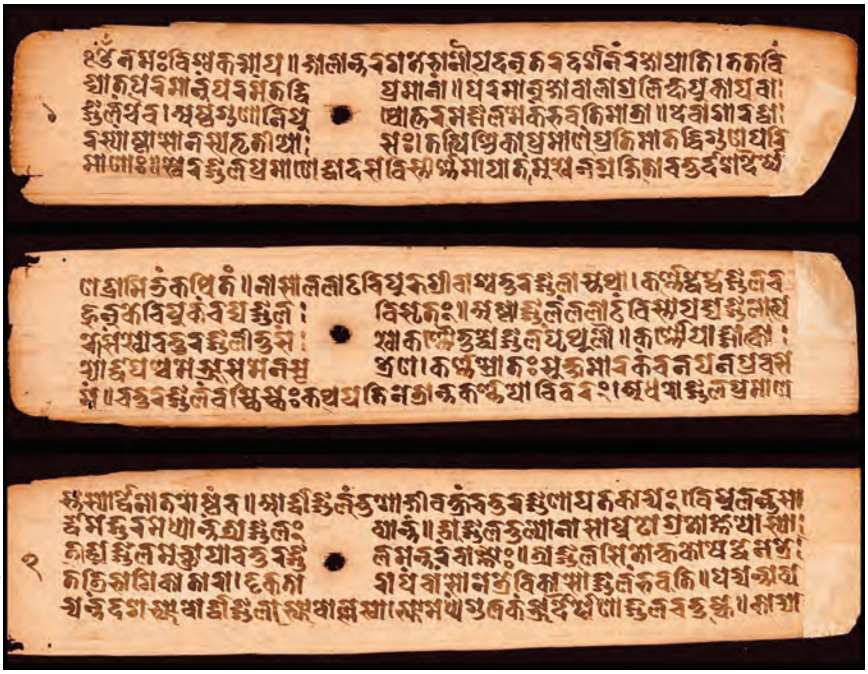


The Hale and Oschin telescopes are among the oldest ones still actively used by researchers. That longevity is a tribute not only to the vision of their builders but also to the efforts of scientists and engineers who over the years built new state-of-the-art spectrographs, cameras, and detectors for the two telescopes. Prominent among those are Gerald Neugebauer and his students, who discovered and studied many of the brightest infrared sources in the sky at Palomar; James Gunn and James Westphal, whose work on CCD detectors at Palomar prepared them to build the *Hubble Space Telescope*'s first-generation camera; and Shrinivas Kulkarni, who led the development of a 600-megapixel camera for the Oschin Telescope. The camera is now being used to detect moving objects, including asteroids and comets, and transient events such as supernovae and gamma-ray bursts.

Schweizer is a research astronomer and science writer, and her experience in both disciplines has resulted in an authoritative book covering a wide breadth of topics. Based on hundreds of interviews she conducted over more than a decade, the book provides a unique historical record of 20th-century astronomy at Palomar. *Cosmic Odyssey* is aimed at the general public, but readers with some background in undergraduate physics will have an easier time with the book.

My only regret is that Schweizer did not include more stories about the colorful personalities who spent their careers at Palomar. She mentions Zwicky's famous comment that some of his colleagues were "spherical bastards"—because they were bastards when viewed from any direction—but that anecdote is only one of many legendary tales from Palomar's long history. I hope that someday she will tell those personal stories too.

Scott Tremaine
University of Toronto
Toronto, Ontario, Canada



ACTIVE IN THE SIXTH CENTURY, the ancient Indian polymath Varāhamihira wrote extensively on topics in astronomy and mathematics. His magnum opus was the encyclopedic text *Bṛhatsaṃhitā* (*The Great Compendium*); this palm-leaf manuscript of that text, held by the Cambridge University Library, dates from 1279 CE.

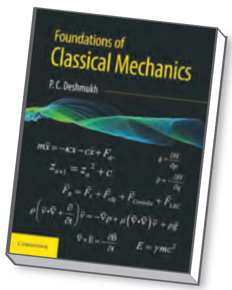
A non-Western take on introductory physics

For most students today, classical mechanics is a stepping-stone on the road to more advanced topics in modern physics or applications in engineering. For that reason, the field may seem unappealing to them. P. C. Deshmukh aims to change that with *Foundations of Classical Mechanics*, a textbook aimed at first-year physics and engineering students. Passionate about the subject, he aims to inspire interest in the field by extracting teachable moments from the exciting discoveries that shaped its history.

To that end, Deshmukh takes every opportunity to connect important lessons from classical mechanics with modern developments in physics. For instance, in a section on symmetry and conservation laws in the first chapter, Deshmukh quotes Albert Einstein's obituary for Emmy Noether and nods to Eugene Wigner's tremendously influential work that made group theory an indispensable part of modern particle physics.

Foundations of Classical Mechanics

P. C. Deshmukh
Cambridge U. Press, 2019. $84.99



Young readers cannot help but be inspired to dig deeply into the foundations of classical mechanics and build a solid base for their careers in physics.

Uniquely, the historical notes have a decidedly Indian perspective. As someone educated in Canada, I was intrigued to learn, among other things, that our modern number system was developed in India; that Varāhamihira studied an earlier version of Pascal's triangle in India in the sixth century; and that the 14th- to 16th-century Kerala school of astronomy and mathematics developed a heliocentric model of the solar system well before the Copernican revolution. It's refreshing and at times a bit shocking to get a glimpse at the history of science from a perspective with less European bias.

As befitting a textbook on classical mechanics, *Foundations of Classical Mechanics* contains chapters on standard material like oscillators, gravity, angular momentum and rotations, and chaos. But the book also includes material on related fields, including two chapters on mathematics and one chapter each on fluid mechanics, electrodynamics, special relativity, and general relativity.

One of the chapters on mathematics is chapter 2, titled "Mathematical Preliminaries." It would have been easy for Deshmukh to bite off more than he could chew in such a chapter, but he wisely chose to limit the scope to coordinate systems and vectors. The latter is taught first from the perspective of Cartesian tensors before more general tensors are introduced. The section's clarity facilitates students' comprehension of later material in the book. It also provides a solid foundation for the presentation of tensor calculus in advanced general relativity textbooks like *A First Course in General Relativity* (2nd ed., 2009) by Bernard F. Schutz or *General Relativity: An Introduction for Physicists* (2006) by M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby.

Special relativity is presented in chapter 13 with 4-vectors, which greatly facilitates the introduction of general relativity in chapter 14. Unfortunately one section of the latter chapter is potentially ambiguous: During a discussion of spacetime intervals, Deshmukh presents readers with a mathematical expression that describes a curved spacetime continuum that is spherically symmetrical. But his wording could leave readers with the impression that the expression is valid for all geometries and not just for that particular case. Still, Deshmukh's presentation of the Lagrangian formulation of Newtonian gravity provides students with a natural and smooth transition to general relativity.

Regarding stylistic matters, I was a bit disappointed to find places with nonstandard or dated language like "man's view of the universe." Fortunately those instances are rare.

Physicists have published tons of textbooks on classical mechanics. Two standard works often assigned at universities are Herbert Goldstein's *Classical*

Mechanics (3rd ed., 2002) and Tom W. B. Kibble and Frank H. Berkshire's *Classical Mechanics* (5th ed., 2004). Goldstein's book is a wonderful resource for advanced physics undergraduates but assigning it to first-year students would throw them into the deep end. The arguments in Kibble and Berkshire's textbook are impressively elegant and rigorous, although I did catch an error in one of the worked examples when I looked at the book recently. Both established books have their strengths, and *Foundations of Classical Mechanics* stands proudly next to those classics.

In his new book, Deshmukh provides a rigorous yet accessible introduction to classical mechanics that is suitable for first- or second-year physics and engineering students. *Foundations of Classical Mechanics* successfully uses a less Western-centric historical perspective to place the field in the context of exciting topics in modern physics.

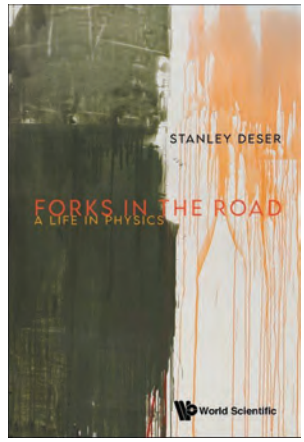
Robert B. Scott
University of Western Brittany
Brest, France

NEW BOOKS & MEDIA

Forks in the Road

A Life in Physics

Stanley Deser
World Scientific, 2021. $58.00



Few physicists alive today can claim to have met such luminaries as Wolfgang Pauli, J. Robert Oppenheimer, and Niels Bohr. Stanley Deser is one of those few. In *Forks in the Road*, the theorist recounts not only his colorful experiences with many such figures in physics but also his own tumultuous life story. Born in 1931 into a Jewish family in Poland, Deser made a harrowing escape with his parents in 1940–41 from Nazi-occupied France to the US via neutral Spain and Portugal. At the end of World War II, he began his study of physics, just as the field started to balloon from a clubby, Old World–based coterie into the massive globe-spanning community it is today. Along with presenting fascinating anecdotes about figures both famous and long forgotten, *Forks in the Road* documents a field's transformation from the inside. –RD

Electrify

An Optimist's Playbook for Our Clean Energy Future

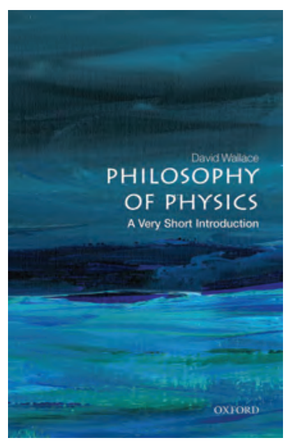
Saul Griffith
MIT Press, 2021. $24.95

As inventor and entrepreneur Saul Griffith puts it in his new book *Electrify*, "It's now time for end-game decarbonization"—namely, to halt the use of fossil fuels immediately. Fortunately, he argues, we have the technology to switch to renewable energy without changing our lifestyles. All we need to do, as the title indicates, is electrify everything, especially cars and heating systems, and build enough renewable energy sources to power it all. Although Griffith largely focuses on the big picture, he also includes helpful advice on how individuals can electrify what he terms our own personal infrastructure. Buying an electric car, installing solar panels on houses, replacing gas- and oil-powered heating systems with electric heat pumps, and choosing energy-efficient electric appliances are all actions individuals can take that significantly reduce carbon emissions. Griffith's refreshingly positive tone undergirds his call to action. –RD

Philosophy of Physics

A Very Short Introduction

David Wallace
Oxford U. Press, 2021. $11.95 (paper)



The latest installment in Oxford University Press's venerable Very Short Introductions series focuses on the philosophy of physics. True to series form, the slim volume authored by philosopher David Wallace presents readers with an overview of the philosophical implications of topics like statistical mechanics, relativity theory, and quantum mechanics. A particular highlight is the first chapter, which, in a breezy 16 pages, covers fundamental questions in philosophy, such as that of scientific realism: Do physical entities like electrons and black holes—which we cannot directly observe—really exist, or are they merely figments of theory that allow us to make predictions? Wallace does an excellent job of presenting all the reasonable philosophical positions on a given topic, even those that he does not personally believe. The book is a superb introduction to a knotty field, and it should appeal both to the educated public and to curious physicists who don't just want to "shut up and calculate." –RD PT

Excellence in Low Temperature Imaging

LT - Scanning Probe Microscope System

KPFM image of $CaFe_2As_2$

Imaging Modes

SHPM, STM, AFM, MFM, EFM
SNOM, Conductive AFM, KPFM,
N- V Center & Confocal Microscopy

Temperature Range

10 mK - Room temperature

Essentially five reasons make researchers adapt their experimental setups to NanoMagnetics Instruments low-temperature system compatibility.

- ☑ Reduced thermal drift
- ☑ Lower noise levels
- ☑ Enhanced stability of tip and sample
- ☑ Reduction in piezo hysteresis/creep
- ☑ Probably the most obvious, the fact that many physical effects are restricted to low temperature

Can be customised to fit in any cryostat

"The LT-AFM/MFM system allows us to perform studies on functional materials to investigate magnetic, piezoelectric and morphological characteristics with nanoscale spatial resolution. The versatility of the system to switch between different measuring modes, and the possibility of working under applied magnetic fields, offers us the possibility to stablish structure-property relationships, fundamental to the understanding, design and use of materials. We are currently applying this technique to the study of vortices dynamics in layered superconductors, and the investigation of ferroelectric/ferromagnetic heterojunctions for spintronic applications."

Dr. Carmen Munuera, 2D Foundry, Material Science Institute of Madrid (ICMM-CSIC)

NanoMagnetics Instruments

/NMInstruments

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

NEW PRODUCTS

Focus on software, instrumentation, and data acquisition

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

Andreas Mandelis

Software for parametric test system

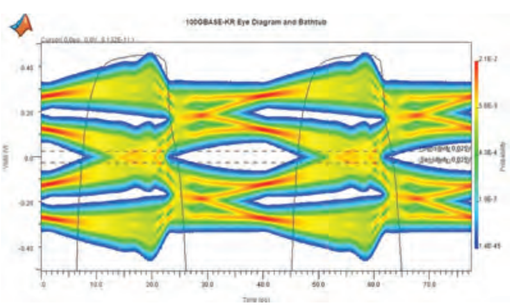
According to Tektronix, its KTE V7.1 software, which is now available for the Keithley S530 Series Parametric Test System, will help accelerate semiconductor-chip manufacturing. The upgraded software and hardware enable high throughput and single-pass testing—users can test up to 1100 V on any pin with a single probe touchdown. Compared with version KTE V5.8, KTE V7.1 reduces test times by more than 10%. Options available for the first time with the KTE V7.1 release include a parallel test capability that optimizes the efficiency of all system resources and a high-voltage capacitance test option for emerging power and wide-bandgap applications. A new test-head design gives users the flexibility to employ a variety of probe cards, and the recently released System Reference Unit shortens calibration time to less than 8 h. ***Tektronix Inc***, *14150 SW Karl Braun Dr, PO Box 500, Beaverton, OR 97077, www.tek.com*

Diaphragm liquid pump

KNF has expanded its series of smooth-flow liquid pumps that combine the low pulsation of gear or centrifugal pumps with the strengths and advantages of diaphragm pumps. The new KNF FP 70 model delivers up to 850 mL/min while producing pressures up to 29.4 psig (2 bar) under continuous operation. Thanks to the patented four-point valves, the KNF FP 70 is reliably self-priming, even at very low motor speeds. It can safely run dry, handles liquid transfer gently and cleanly, and is available with chemically resistant flow-path materials. Integrated pulsation-dampening technology eliminates the need for additional pulsation-dampening elements and tubing. The versatile FP 70 is available with a selection of motors from high-end brushless DC to lower-end DC. Other options include a selection of hydraulic connections and various flow-path materials. Applications for the FP 70 include analytical instruments, medical technology, and 3D printing. ***KNF Neuberger Inc***, *2 Black Forest Rd, Trenton, NJ 08691-1810, https://knf.com*

Software-defined instrumentation platform

The Moku:Pro system from Liquid Instruments is a software-upgradeable platform that integrates multiple standalone devices and allows users to dynamically switch between them. Based on the field-programmable gate array, the Moku:Pro hosts nine instruments with real-time measurement capabilities: an oscilloscope, DC–600 MHz lock-in amplifier, proportional-integral-derivative controller, phasemeter, arbitrary waveform generator, data logger, spectrum analyzer, frequency response analyzer, and waveform generator. Advanced analog-to-digital converter blending technology ensures that each instrument functions with optimal sensitivity from RF to acoustic frequencies. Though designed to support research in physics and engineering in fields such as aerospace and semiconductors, Moku:Pro's instrument suite is particularly suitable for photonics applications, including spectroscopy, microscopy, metrology, gravitational-wave detection, active laser stabilization, and quantum computing. ***Liquid Instruments***, *740 Lomas Santa Fe Dr, Ste 102, Solana Beach, CA 92075, www.liquidinstruments.com*

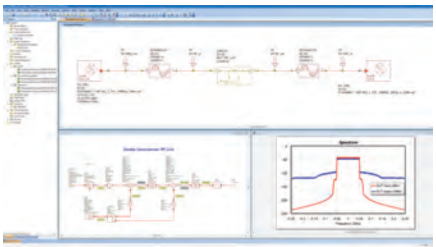


Mathematical programming software

MathWorks is now offering Release 2021b (R2021b) of its MATLAB and Simulink software. In addition to many updated features and functions, new capabilities in MATLAB include code refactoring, block editing, and the ability to run Python commands and scripts from MATLAB. Simulink updates let users run multiple simulations for different scenarios from the Simulink Editor and create custom tabs in the Simulink Toolstrip. R2021b introduces two new products that support wireless communications: The RF PCB Toolbox enables the design, analysis, and visualization of high-speed and RF multilayered printed circuit boards; and the Signal Integrity Toolbox provides functions and apps for designing high-speed serial and parallel links. R2021b also includes major updates to the Symbolic Math and Lidar Toolboxes, Simulink Control Design, and other products in the areas of predictive maintenance, deep and reinforcement learning, and statistics and machine learning. ***The MathWorks Inc***, *1 Apple Hill Dr, Natick, MA 01760-2098, www.mathworks.com*

RAN analytics software

Nemo 5G RAN Analytics software from Keysight is a fully automated cloud-based solution designed to speed analysis of a mobile operator's 5G radio access network (RAN) performance and to streamline data processing, reporting, and analytics. Based on a centralized, web-based data management platform, the software combines data analytics built on artificial intelligence and machine learning frameworks with an intuitive user interface to efficiently manage large amounts of data captured in a live 5G network. Nemo 5G RAN Analytics is available as software as a service. It connects with Keysight's other Nemo solutions, including the Outdoor 5G NR Drive Test Solution, a laptop-based software for measuring the real quality of the end user's experience; the Cloud Remote Monitoring Solution; and the Handy handheld measurement software. When combined, Keysight's Nemo solutions enable users to quickly and reliably upload data captured in the field and share a common set of analytics reports across the organization. ***Keysight Technologies Inc***, *1400 Fountaingrove Pkwy, Santa Rosa, CA 95403-1738, www.keysight.com*



Software for signal creation and analysis

Rohde & Schwarz and Cadence have collaborated to produce a tool for signal creation and analysis to help speed the development of RF components used in wireless communications and radar design. The R&S VSESIM-VSS functions as an addition to the Cadence Visual System Simulator (VSS) software. It expands the capabilities of the VSS software by adding realistic signals for both simulation and testing, thereby increasing simulation accuracy and simplifying the design process. The tool combines the signal-generation and signal-analysis functions from two R&S software tools for testing operative circuits, modules, and devices—the WinIQSIM2 simulation software and the VSE vector signal explorer—and adds plug-ins for Cadence electronic design automation tools. The data sink plug-ins provide access to the signal at any point in the design process. The R&S VSESIM-VSS supports all major standards such as 5G and the latest Wi-Fi evolutions. ***Rohde & Schwarz GmbH & Co KG***, *Muehldorfstrasse 15, 81671 Munich, Germany, www.rohde-schwarz.com*

Motion-control platform

Aerotech has upgraded its Automation1 motion-control platform, adding several new drives and other improvements to version 1.2. Linear-Theta gantry control, which is now included in the Software-Based Machine Controller (iSMC), enables more flexibility and better performance for gantries. The iSMC's new C language Application Programming Interface library lets users easily integrate any or all Automation1 functions within their custom applications. CADFusion has been updated to export to Automation1's AeroScript programming language format, which lowers implementation risk and delivers improvements in part quality and throughput. AutoFocus, designed to simplify sensor integration and increase system performance in surface profiling and scanning applications, is now standard on all Automation1 drives. Also standard is the Enhanced Throughput Module, which improves the positioning performance of high-dynamic motion systems by directly measuring the unwanted motion of the machine base and communicating it back to the controller. ***Aerotech Inc***, *101 Zeta Dr, Pittsburgh, PA 15238-2811, www.aerotech.com*

High-speed waveform-generator-and-digitizer platform

Spectrum Instrumentation has added eight high-speed models to its range of hybridNETBOX products. The hybridNETBOX instrumentation platform combines a multichannel arbitrary waveform generator (AWG) and a digitizer in a single portable unit. Capable of simultaneous signal generation and acquisition, the instruments are suitable for applications involving stimulus-response or closed-loop-type testing. The new models offer a choice of two or four AWG channels combined with the same number of digitizer channels. The AWG channels can generate almost any waveshape. Models are available with output rates of either 625 MS/s or 1.25 GS/s and signal bandwidths up to 400 MHz, with 600 MHz optional. For signal acquisition, the digitizer channels provide 16-bit resolution and sampling rates of 180 MS/s or 250 MS/s, or 14-bit resolution and sampling rates of 400 MS/s or 500 MS/s. Both the AWG and the digitizer feature a flexible clocking system that allows users to select almost any output or sampling-rate setting, which enables them to generate or acquire signals at exactly the required speed. ***Spectrum Instrumentation Corp***, *401 Hackensack Ave, 4th Flr, Hackensack, NJ 07601, https://spectrum-instrumentation.com*

Thermal analysis software

Waters Corporation has made available TRIOS AutoPilot software for use with the thermal-analyzer product line of its TA Instruments division. The software is designed to help materials science laboratories conducting R&D and quality control to automate and standardize workflows. By simplifying the process of making analytical measurements, the software can improve laboratory efficiency, ensure product quality, and reduce risks, according to the company. TRIOS AutoPilot is the first thermal analysis software to be based on Google's visual programming interface, Blockly. With that intuitive open-source software, users can create custom scripts and configure them for thermal-analysis applications without having to learn a higher-level programming language. The software comes with prewritten express scripts for automating many common procedures such as sample loading, instrument calibration and verification, and formatting and exporting data files into laboratory information management systems. The OneTouch interface guides operators with video and text prompts. ***Waters Corporation****, 34 Maple St, Milford, MA 01757, www.waters.com*

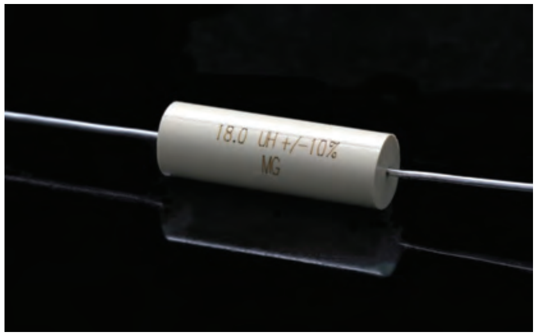
Data-acquisition software

CAS DataLoggers has introduced the ProfiSignal 20 (PS20) software package from Delphin Technology. The software allows data collected with Delphin data-acquisition systems to be visualized and analyzed on a personal computer, tablet, or smartphone. The Go version of PS20 can be used for quickly visualizing, monitoring, and archiving measurement data. It includes configurable single and multiaxis trend charts and multitrack diagrams that allow measurement data from multiple channels, both digital and analog, to be displayed synchronously; statistical functions for data analysis; and data export for third-party software applications. The Basic version of PS20 extends the capabilities of Go with a range of configurable control and monitoring widgets that can be used to create user-defined dashboards. A key innovation in PS20, the Scan and Check function, allows projects to be built with a custom QR code. When a smartphone scans the code, it displays live data for the project, including any custom visualizations. ***CAS DataLoggers****, 8437 Mayfield Rd, Unit 104, Chesterland, OH 44026, www.dataloggerinc.com*

Hydraulic data logger

Webtec has unveiled the versatile HPM7000 handheld hydraulic data logger, the successor to its long-standing HPM6000 model. The HPM7000 lets users working in R&D or doing testing, repairs, or inspections connect up to 48 sensors and create different hardware configurations simply by adding or changing input modules. Test-stand technicians can monitor and record variables such as flow, temperature, pressure, current, rotational speed, and contamination, for example, following the repair or rebuilding of pumps, motors, pistons, or valves. R&D users developing and analyzing hydraulic machinery such as tractors, excavators, and industrial systems can use the HPM7000 to collect traditional flow, temperature, and pressure readings and seamlessly integrate new J1939 or CANOpen sensors of different types or manufacturers into a single data-logging device. The compact data logger can be powered by a battery or an external power supply and used as a portable or fixed device. ***Webtec LLC***, *1290 E Waterford Ave, St. Francis, WI 53235, https://en.webtec.com*

Nonmagnetic RF inductors



Gowanda Electronics designed its 28MG series of nonmagnetic RF through-hole (leaded) inductors to address the need to achieve inductance values of up to 18 μH. The series is suitable for use in MRI and specific types of x-ray equipment and other applications where the presence of magnetic materials could compromise system performance. Those include telecommunications; instrumentation; and equipment for laboratory analysis, electronic testing, aviation, and navigation. The performance range provided by the 60 discrete parts within the 28MG series includes inductance from 1.2 μH to 18 μH, DC resistance from 0.079 Ω to 4.15 Ω, and DC current rating from 315 mA to 2400 mA. Gowanda's nonmagnetic surface mount and through-hole inductors provide relative permeability of less than or equal to 1.00003. The 28MG series' operating temperature range is −55 °C to 125 °C. The 28MG inductors are epoxy encapsulated to withstand all types of reflow soldering and protect the environment. ***Gowanda Electronics***, *1 Magnetics Pkwy, Gowanda, NY 14070, www.gowanda.com* PT

OBITUARIES

Myriam Paula Sarachik

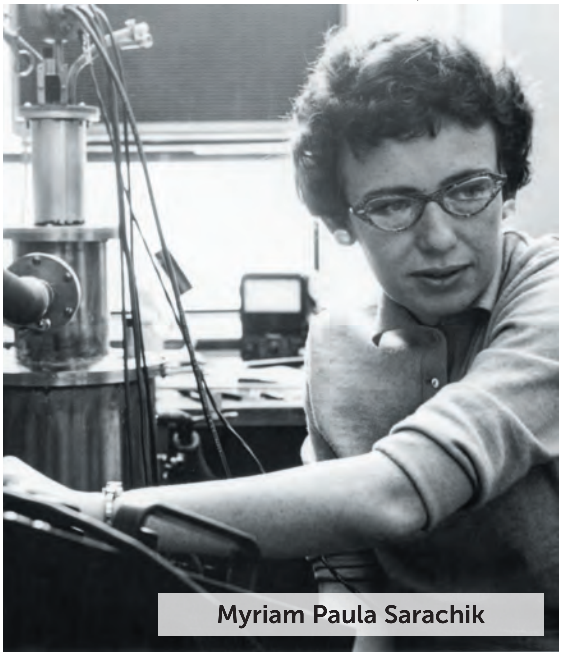
On 7 October 2021, the physics community lost a towering figure, Myriam Sarachik. Over six decades she made outstanding contributions to experimental condensed-matter physics.

Myriam Paula Morgenstein was born in Antwerp, Belgium, on 8 August 1933. She was in the first grade when the Nazis invaded the country. Part of her extended Jewish family perished in a gas chamber. Myriam, her younger brother Henry, and their parents managed to escape through occupied France to Spain. In 1941 they arrived in Cuba, which granted them refugee status. Myriam took years of piano lessons in Havana and was so good that she dreamed of becoming a concert pianist. In 1947 the family immigrated to New York. Myriam attended the Bronx High School of Science, where her class of 1950 included Steven Weinberg and Sheldon Glashow.

When she entered Barnard College, music was still Myriam's passion, but her interest turned to physics. She took advanced physics classes at Columbia University, where she met an engineering student, Philip Sarachik, whom she shared her life with until her death. She received her bachelor's degree in physics in 1954. After working at IBM's Watson Scientific Computing Laboratory at Columbia, Myriam earned her master's and PhD from the university in 1957 and 1960, respectively. Her doctoral work, under the supervision of Richard Garwin, provided one of the first demonstrations of the Bardeen-Cooper-Schrieffer superconducting gap.

Myriam's bosses were skeptical about her pursuing a career in a field dominated by men. At job interviews, she was advised to stay home and take care of her daughter, Karen, who was born soon after Myriam received her PhD. She finally managed to secure a research position at Bell Labs. Her work there provided in 1964 the first experimental evidence of the Kondo effect—the minimum of metallic resistivity as a function of temperature in the presence of magnetic impurities. Myriam did all the measurements herself and wrote the article. Her coauthors contributed by making the samples. That year she joined the physics faculty of the City College of New York, where she spent the next 54 years.

AIP ESVA/GIFT OF M. SARACHIK



Myriam Paula Sarachik

In 1970 tragedy struck. Myriam and Phil's five-year-old second daughter, Leah, was abducted. Her body was found in Vermont many weeks later. Myriam was eventually able to return to teaching, but it took her more than a decade to restart research.

In the early 1980s, Myriam and her students began studies of the effect of dopants on the magnetoresistance of various materials. They elucidated quantum interference phenomena in electron scattering and found evidence of the transition from Mott to Efros–Shklovskii variable-range hopping as the temperature is lowered. She then turned her attention to the metal–insulator transition in two dimensions. Her work in the 1990s with postdoc Sergey Kravchenko resulted in discoveries of the universal scaling of 2D resistivity with the electric field and of the suppression of metallic conductivity by an in-plane magnetic field.

During that time, Roberta Sessoli in Italy brought a new system to the attention of physicists—crystallites of weakly coupled magnetic molecules of spin-10. Myriam was interested in the possibility of quantum flipping of the molecules' magnetic poles. She put graduate student Jonathan Friedman on the project and also invited the University of Barcelona's Javier Tejada, who had experience with measurements of magnetic particles, to join her. Their initial results were interesting but inconclusive. Many people would have published them and moved on, but that was not Myriam's way. Her insistence on clarity and the brilliance of the team she assembled eventually led to a breakthrough. In 1996 they unambiguously demonstrated quantum tunneling of the magnetic moment by a spectacular stepwise magnetization curve of manganese-12 acetate.

That work brought molecular magnetism out of infancy into the field of quantum information technology, and it contributed to Myriam earning the Oliver E. Buckley Condensed Matter Physics Prize of the American Physical Society (APS). She and members of her lab followed that work by discovering in 2005 the phenomenon of magnetic deflagration, in which the Zeeman energy of a molecular magnet burns in a manner similar to flame propagation.

Despite tragedies in her life, Myriam was a kindhearted and caring person. She was passionate about bringing more women into physics, a vision she promoted when serving as APS president in 2003. Going to dinners with her and her students was great fun. She would never refuse a glass of wine. She had encyclopedic knowledge of things far beyond physics. Discussions with her were always interesting and memorable. She retired from City College in 2018. Two years later she was awarded APS's highest prize, the APS Medal for Exceptional Achievement in Research. Even as she was becoming frail, her mind remained clear and fast until the end. Her death has left a large hole in the hearts of her colleagues, students, and friends.

Eugene Chudnovsky
City University of New York
New York City PT

QUICK STUDY

Kostya Trachenko is a professor of physics at Queen Mary University of London in the UK. **Vadim Brazhkin** directs the Institute for High Pressure Physics in Moscow.

The quantum mechanics of viscosity

Kostya Trachenko and Vadim V. Brazhkin

Although a liquid's viscosity depends strongly on temperature and pressure, its minimum value is fixed by fundamental physical constants.

In the 1970s, physicist and Nobel laureate Edward Purcell noticed that no liquid exists with a viscosity much lower than that of water. In the first paragraph of his essay "Life at low Reynolds number," he writes, "Viscosities have a big range *but they stop at the same place*. I don't understand that."

By stopping "at the same place," Purcell meant that liquid viscosities never drop below a certain value. And he says in the first footnote of the essay that Victor Weisskopf had explained the phenomenon to him. To date, however, no one has ever found a published account of the explanation. Even so, Weisskopf published his own essay "About liquids" around the time Purcell published his. That paper starts with a sobering story of the challenges that theoretical physicists face when they try to deduce the states of matter using only quantum mechanics. They can predict the existence of gases and solids, but not liquids.

The upshot is that liquids are difficult—a point hammered home in textbooks. Lev Landau and Evgeny Lifshitz's *Statistical Physics*, for instance, repeatedly asserts that the thermodynamic properties and temperature dependences of a liquid simply cannot be calculated in analytic form applicable to all liquids. The reason is a combination of strong molecular interactions and the absence of small oscillations that simplify the theory of solids. That complication is embodied in the famous "no small parameter" problem: Liquids have neither the weak interactions of a gas nor the small atomic displacements of a solid. Despite the difficulty, we have developed our own theory of liquid thermodynamics, based on excitations in liquids, which is currently undergoing detailed tests.

Viscosity minimum

Meanwhile, we can ask whether theorists understand viscosity well enough to answer Purcell's question of why all viscosities stop at the same place. Viscosity η denotes a liquid's resistance to a shear force and governs important properties such as diffusion and dissipation. In a dilute gas-like fluid, η is set by molecules moving at distances up to the mean free path L and transferring momentum during collisions: Specifically, $\eta = \rho v L/3$, where ρ and v are the density and average velocity of molecules, respectively.

That equation predicts that the viscosity of a gas increases with temperature, because molecular velocity increases with temperature. That prediction is counterintuitive because fluids usually thin when they are heated. Unlike gases, dense liquids have a viscosity set by their molecules vibrating around quasi-equilibrium positions before jumping to neighboring sites. The frequency of those jumps increases with temperature, and viscosity consequently decreases with temperature: $\eta = \eta_0\exp(U/k_BT)$, where U is the activation energy.

The increase of viscosity at high temperature and its decrease at low temperature imply that it has a minimum. That minimum arises from the crossover between two different viscosity regimes: a gas-like regime, where the kinetic energy of higher-temperature particles provides a larger momentum transfer, and hence larger η, and a liquid-like regime, where lower temperature decreases the frequency of site-jumping particles and slows down the liquid flow, also resulting in larger η.

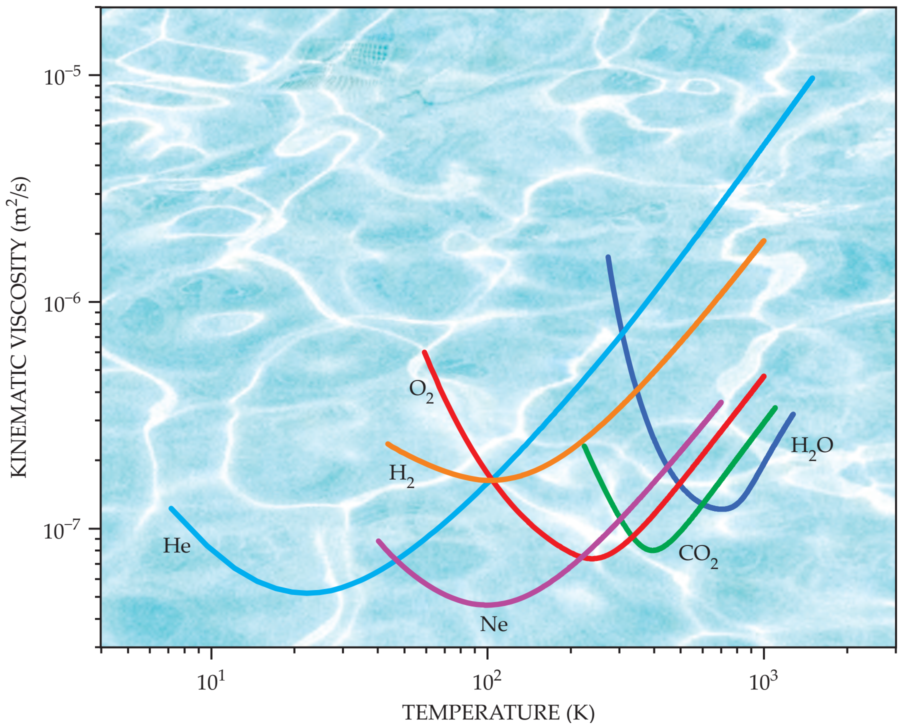
It is convenient to look at that crossover above the critical point where it is smooth and no liquid–gas phase transition intervenes. Think in terms of kinematic viscosity $\nu = \eta/\rho$, which describes the properties of liquid flow. The figure on page 67 shows experimental values of several supercritical fluids. Kinematic viscosities clearly have minima, and one can understand them as the crossover states between gas-like and liquid-like behavior.

The viscosity minima provide the first clue to the Purcell question: Viscosities stop decreasing, of course, when they hit their minima. But can each minimum itself get arbitrarily close to zero? (Note that we do not discuss superfluidity in this Quick Study.) Why are the minima of η hard to move up or down and somehow close to the viscosity of water at ambient conditions?

Scientists could answer that question if they could calculate the viscosity at its minimum. But that's complicated, as Landau and Lifshitz discuss in their book. Molecular interactions are strong and system specific. Calculating viscosity parameters even in simple liquids is difficult using only theory and no input from modeling. And for molecular liquids, such as water, it's nearly impossible.

A revealing approximation

Fortunately, the minimum of viscosity at the crossover is a special point where viscosity can, in fact, be evaluated, if only approximately. The minimum values ν_{min} are related to just two basic properties of a condensed-matter system $\nu_{min} = \omega_D a^2/2\pi$, where a is the interatomic separation and ω_D is the system's Debye frequency. The two parameters can, in turn, be related to the radius of the hydrogen atom and a characteristic



EXPERIMENTAL KINEMATIC VISCOSITIES of noble and molecular liquids. Each one exhibits a minimum. Viscosities of helium, hydrogen, oxygen, neon, carbon dioxide, and water are plotted at 20 MPa, 50 MPa, 30 MPa, 50 MPa, 30 MPa, and 100 MPa, respectively. (Source: NIST, https://webbook.nist.gov/chemistry/fluid.)

bonding strength set by the Rydberg energy. Then ν_{min} becomes

$$\nu_{min} = \frac{1}{4\pi}\frac{\hbar}{\sqrt{m_e m}}, \quad (1)$$

where m_e is the mass of an electron and m is the mass of the molecule.

Two fundamental constants $\hbar$ and m_e appear in that equation. The minimal viscosity turns out to be quantum! That may seem surprising and at odds with our concept of high-temperature liquids as classical systems. But equation 1 reminds us that the nature of interactions in condensed matter is quantum mechanical, with $\hbar$ affecting both the Bohr radius and Rydberg energy.

The fundamental constants help keep ν_{min} from moving up or down much. And because ν_{min} is inversely proportional to the square root of the molecule's mass, the viscosities themselves are not universal—although that does not change ν_{min} much. For different liquids, such as those plotted in the figure, equation 1 predicts that ν_{min} should fall in the range $(0.3–1.5) \times 10^{-7}$ m²/s. That range is reassuringly close to experimental values.

Therefore, the answer to the Purcell question is that viscosities stop decreasing because they have minima, and those minima are fixed by fundamental constants. Interestingly, the same happens to an unrelated liquid property, thermal diffusivity, which governs how well liquids transfer heat. That variable also exhibits minima given by equation 1. It does so because thermal diffusivity depends on the same two parameters as ν_{min}, a and ω_D.

As shown in equation 2, when m is set equal to the proton mass m_p, equation 1 gives rise to a universal quantity ν_f, the fundamental kinematic viscosity:

$$\nu_f = \frac{1}{4\pi}\frac{\hbar}{\sqrt{m_e m_p}} \approx 10^{-7}\ \text{m}^2/\text{s}. \quad (2)$$

The fundamental physical constants $\hbar$, m_e, and m_p are of general importance. Together with the electron charge and speed of light, they form dimensionless constants that determine whether the universe is biofriendly. That's because they affect the formation of stars and the synthesis of heavier elements, including carbon, oxygen, and so on, which can then form molecular structures essential to life.

Fundamental constants and water

The fundamental constants are friendly to life at a higher level too: Biological processes, such as what happens in cells, rely heavily on water. Were Planck's constant to take on a different value, for instance, water viscosity would change too—both its kinematic viscosity ν, which is relevant to water flow, and its dynamic viscosity η, which sets internal friction and diffusion. If the viscosity minimum were to increase because of a higher value of $\hbar$, for instance, water would become more viscous, and biological processes would not be the same. Life might not exist in its current form or even at all.

One might hope that cells could still survive in such a universe by finding a hotter place where overly viscous water becomes thinned. That would not help, though. Planck's constant sets the minimum below which viscosity cannot fall, regardless of temperature. Water and life are indeed well attuned to the degree of quantumness of the physical world.

One can hope that Purcell would have been happy with the answer to his question. Unless he already heard it from Weisskopf in the 1970s.

Additional resources

▶ E. M. Purcell, "Life at low Reynolds number," *Am. J. Phys.* **45**, 3 (1977).

▶ V. F. Weisskopf, "About liquids," *Trans. N. Y. Acad. Sci.* **38**, 202 (1977).

▶ L. D. Landau, E. M. Lifshitz, *Statisticheskaia fizica* (*Statistical Physics*), 2nd ed., Pergamon Press (1969).

▶ J. E. Proctor, "Modeling of liquid internal energy and heat capacity over a wide pressure–temperature range from first principles," *Phys. Fluids* **32**, 107105 (2020).

▶ K. Trachenko, V. V. Brazhkin, "Minimal quantum viscosity from fundamental physical constants," *Sci. Adv.* **6**, eaba3747 (2020).

▶ J. D. Barrow, *The Constants of Nature: From Alpha to Omega—The Numbers That Encode the Deepest Secrets of the Universe*, Pantheon Books (2003). PT

Magdeburg hemispheres

A Magdeburg sphere consists of a pair of hollow hemispheres, historically made from copper or brass, with a seal around the edge that's tight enough to allow air to be pumped out of the assembled object. When air is evacuated from the sphere, the pressure inside of it is lower than the atmospheric pressure outside pushing on it from all directions, so the halves stay together. The inventor, Otto von Guericke, used such a sphere to perform a dramatic demonstration of the existence of atmospheric pressure. During the experiment, air was pumped out of the sphere, and a team of horses was attached to either side. They strained in opposite directions yet were unable to pull the hemispheres apart until air was let back in.

The demonstration was not only a scientific endeavor but a political move. Von Guericke, then mayor of the city of Magdeburg, in what is now Germany, held the demonstration during the Holy Roman Empire's diet at Regensburg in 1654 to show the city's power and its recovery. In 1631 during the Thirty Years' War, 20 000 inhabitants were slaughtered by Imperial forces in what became known as the Sack of Magdeburg. Von Guericke attracted the attention of Emperor Ferdinand III, who ordered Gaspar Schott, a mathematics teacher in Würzburg, to perform more such experiments using von Guericke's equipment. In Schott's first published book on vacuum science, *Mechanica hydraulico-pneumatica* (*Hydraulic-Pneumatic Mechanics*, 1657), he promoted von Guericke's research and instruments. This illustration was created by Schott for a large foldout page in von Guericke's major work, *Experimenta nova (ut vocantur) Magdeburgica de vacuo spatio* (*The New [So-Called] Magdeburg Experiments of Otto von Guericke*, 1672). A copy is available to view at the Niels Bohr Library and Archives. —CEM

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

Semantic segmentation for wildlife conservation.