Newsjournal of the Society for Industrial and Applied Mathematics
sinews.siam.org

siam news

Volume 50/ Number 1
January/February 2017

Data-driven Discovery of Governing Physical Laws

Dynamical Systems and Machine Learning

By Steven L. Brunton, J. Nathan Kutz, and Joshua L. Proctor

Ordinary and partial differential equations are widely used throughout the engineering, physical, and biological sciences to describe the physical laws underlying a given system of interest. We implicitly assume that the governing equations are known and justified by first principles, such as conservation of mass or momentum and/or empirical observations. From the Schrödinger equation of quantum mechanics to Maxwell's equations for electromagnetic propagation, knowledge of the governing laws has allowed transformative technology (e.g., smart phones, internet, lasers, and satellites) to impact society. In modern applications such as neuroscience, epidemiology, and climate science, the governing equations are only partially known and exhibit strongly nonlinear multiscale dynamics that are difficult to model. Scientific computing methods provide an enabling framework for characterizing such systems, and the SIAM community has historically made some of the most important contributions to simulation-based sciences, including extensive developments in finite-difference, finite-element, spectral, and reduced-order modeling methods.

The plummeting cost of sensors, computational power, and data storage in the last decade has enabled the emergence of data methods for the sciences. Such vast quantities of data offer new opportunities for data-driven discovery, referred to as the fourth paradigm of science [7]. Of course, data science is not new. More than 50 years ago, John Tukey envisioned the existence of a scientific effort focused on learning from data, or *data analysis* [5, 16]. Eventually two cultures, centered on the concepts of *machine learning* and *statistical learning*, emerged within the community of data scientists [1]. The former focuses on prediction, while the latter concerns inference of interpretable models from data. Both methodologies have achieved significant success across many areas of big data analytics. But these traditional approaches fall short of achieving a general goal for computationally-oriented scientists, which is inferring a (typically nonlinear) model from observations that both correctly identifies the underlying dynamics *and* generalizes qualitatively and quantitatively to unmeasured parts of phase, parameter, or application space.

Nowhere are these philosophical outlooks more clearly illustrated than in the historical developments concerning planetary motion and gravitation by Johannes Kepler (1571-1630) and Sir Isaac Newton (1643-1727). Both were leading figures of the scientific revolution, which many consider to have begun with Nicolaus Copernicus's *De revolutionibus orbium coelestium* (*On the Revolutions of the Heavenly Spheres*). This work displaced the Ptolemaic doctrine of the perfect circle that had been the dominant predictive theory for nearly 1,500 years.

Kepler was an early big data scientist. As an assistant to Tycho Brahe, he had access to the best and most well-guarded astronomical data collected to date. Upon Brahe's untimely death, Kepler was appointed his successor with the responsibility to complete Brahe's unfinished work. Over the next eleven years, he laid the foundations for the laws of planetary motion, positing

See **Data-driven Discovery** *on page 4*

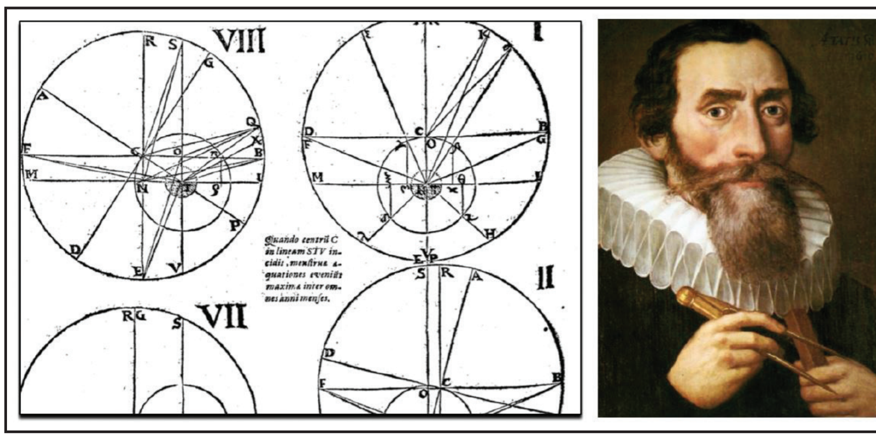


Figure 1. *Using Tycho Brahe's state-of-the-art data, Johannes Kepler utilized geometrical principles in Tabulae Rudolphinae [8] to discover that planetary orbits were actually ellipses. Figure credit: [8] (left) and Creative Commons (right).*

Advantage of Diversity: Consensus Because of (not Despite) Differences

By Takashi Nishikawa and Adilson E. Motter

One generally assumes that individual entities are more likely to exhibit the same behavior if they are equal to each other – imagine animals using the same gait, lasers pulsing together, birds singing the same notes, and agents reaching consensus. In a recent study [2], we demonstrated that this assumption is in fact false for networks of coupled entities. The behavior underlying this finding is an instance of a new network phenomenon we dubbed asymmetry-induced symmetry (AIS), in which the state of the system can be symmetric only when the system itself is not.

We consider spontaneous synchronization in a network of n identically-coupled oscillators as a convenient model process to illustrate the core idea of AIS. In this process, the oscillators synchronize by reaching a stable state in which they all exhibit the exact same dynamics: $x_1(t)=x_2(t)=\ldots=x_n(t)$ for all t. The state of the network then has maximum symmetry, since any two nodes can be swapped without changing the state. It might be intuitive to assume that complete synchronization would require that the oscillators themselves be identical, or at least it would be observed for identical oscillators if it were possible for non-identical ones. The rationale for this is that if the oscillators have identical coupling patterns, complete synchronization of the entire network is a state inheriting the symmetry of the system only if all of the oscillators are identical. The possibility of AIS shows, however, that scenarios exist in which all oscillators synchronize and have identical states if and only if the oscillators themselves are not identical. For the model system illustrated in Figure 1, this remarkable behavior is generic when the coupling is directional. This behavior is also prevalent when the oscillators are not identically coupled, although it is more interesting to first consider the identically-coupled case, in which the need for non-identical oscillators is not only likely but certain to break the symmetry of the system to preserve the symmetry (and stability) of the state.

AIS can be interpreted as the converse of the well-studied phenomenon of symmetry breaking, where the state has less symmetry than the system. Symmetry breaking underlies, for example, the phenomenon of superconductivity, the mechanism through which some elementary particles have mass, and various patterns of network dynamics; it also underlies previously studied (divergent) forms of pattern formation, in which initially symmetric structures evolve into asymmetric ones. While we use synchronization to illustrate AIS because synchronization has long served as a paradigm for emergent behavioral uniformity, the phenomenon has far-reaching implications for any process that involves converging to uniform states. For example, it offers a mechanism for convergent forms of pattern formation in which an asymmetric structure develops into a symmetric one, such as in the development of fivefold radial symmetry in adult starfish from bilateral symmetry in starfish larvae. AIS also has implications for consensus dynamics, potentially yielding scenarios in which interacting agents only reach consensus when they are sufficiently different from each other; this means that diversity may facilitate, and even be required for, consensus.

It is instructive to interpret this phenomenon in the context of Curie's principle [1], which asserts that the symmetries of the causes must be found in the effects. AIS requires that (i) any state with the symmetry of the system be unstable and (ii) the symmetry of the system be reduced to stabilize

See **Diversity** *on page 3*

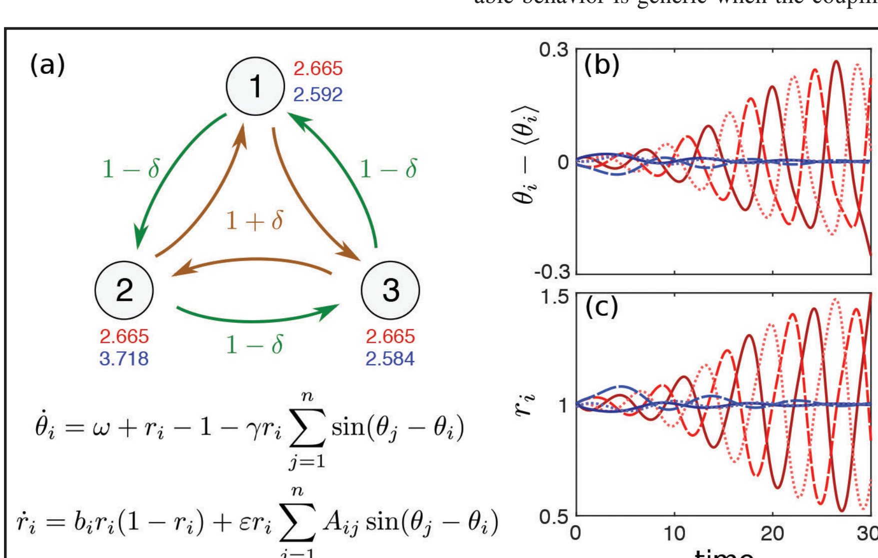


Figure 1. *Three-oscillator network exhibiting AIS.* **1a.** *Structure of the network and the equation of motion. The red and blue numbers are values of the parameters b_i in the case of identical and non-identical oscillators, respectively. The other parameters are $\omega=1$, $\gamma=0.65$, and $\varepsilon=2$.* **1b and 1c**. *Angles θ_i (**b**) and amplitudes r_i (**c**) as functions of time for identical (red) and non-identical (blue) oscillators, showing unstable and stable synchronization, respectively. The notation $\langle\cdot\rangle$ indicates average over i. Image credit: Takashi Nishikawa and Adilson E. Motter.*

Nonprofit Org
U.S. Postage
PAID
Permit No 360
Bellmawr, NJ

siam
SOCIETY for INDUSTRIAL and APPLIED MATHEMATICS
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 USA

siam news
Volume 50/ Number 1/ January/February 2017

Evolving and Innovating

FROM THE SIAM PRESIDENT

By Nicholas Higham

I start my term as SIAM President 33 years after I first joined SIAM as a student member. That year, 1984, I attended the SIAM Summer Meeting—the equivalent at the time of the SIAM Annual Meeting—in Seattle, WA. It had 692 attendees and included 10 invited presentations and 20 minisymposia, delivered with a parallelism of four.

By comparison, the 2016 SIAM Annual Meeting in Boston, MA, had two and a half times as many attendees, 60% more invited speakers, and nearly eight times as many minisymposia, with a parallelism of up to 17 – and that's excluding the many activities at the SIAM Conference on the Life Sciences, with which the meeting was jointly held. SIAM conferences have come a long way in 33 years.

The strapline for the 1984 meeting read "A week-long SIAM meeting featuring topics in applied mathematics of broad interest, plus a trip to the Northwest with opportunities for boating, hiking, camping, bicycling, fishing, tennis and good eating." This tells us two things: SIAM did not always use the Oxford comma, and in those days attendees had time to think about leisure pursuits. (In case you're wondering how I can remember these details, the booklet for the meeting is available under "Archives and Future Meetings" on the Conferences page of the SIAM website, and I recommend browsing the archives as a way to relive memories of past conferences).

This is an exciting time for SIAM, with changes happening that affect all aspects of SIAM's operations, and I look forward to helping address the opportunities that lie ahead in the next two years. Here, I will outline some of the things that are on my mind, and I welcome feedback from readers.

For many years, SIAM leadership has been aware that publishing models are changing, and that SIAM's journal programme must be prepared to adapt in response to these changes. SIAM has continued to introduce new journals, most recently the *SIAM Journal on Applied Algebra and Geometry* (electronic only, like all new SIAM journals), and has catered to open-access publication by allowing authors to pay an article processing charge (currently $2,500) to make a paper freely downloadable. Nevertheless, because journal subscriptions have remained steady, SIAM has not yet made any major changes to its publishing model. My knowledge of new journals from other publishers, such as *Forum of Mathematics* and *PeerJ Computer Science*, on whose editorial boards I sit, has taught me that almost every aspect of SIAM journals could be handled differently. I am keen for us to experiment with some changes. A failed experiment could be richly compensated by a successful one that moves us ahead of the competition or introduces a new development that would be inevitable in a few years' time. We also need to keep abreast of the latest technologies, such as Open Researcher and Contributor ID (ORCID), which is scheduled for implementation in the SIAM journal submission process by early 2017, and the evolving scholarly publishing tools provided by the CrossRef organization.

Returning to the topic of conferences, less than 25% of SIAM members attend a SIAM conference in any given year. For members who live outside North America, attending a conference may be inconvenient or prohibitively expensive. I am interested in whether we can do more to allow those unable to attend a particular SIAM conference to benefit from it. We currently make slides and audio of selected lectures available on SIAM Presents, and publish articles on notable talks in *SIAM News*. Each meeting also has its own hashtag to encourage tweeting. Is there more we should be doing?

Conference attendees enjoying lunch during the 1997 SIAM Annual Meeting at Stanford University. SIAM photo.

SIAM has a well-developed "people pipeline," which allows volunteers to work their way up from, say, being a student chapter officer to an activity group (SIAG) officer or a member of one of SIAM's many committees, perhaps going on to become an elected council or board member or to hold one of the major offices. This process works only if SIAM identifies and encourages suitable people, particularly in a way that reflects membership diversity. This is not an easy task, and the Committee on Committees and Appointments (which does not appoint to itself!) works very hard each year to produce a suitable set of appointments, taking into account names suggested via recommendations[1] on the Officers, Board, and Council page of the SIAM website. Not only do we need to appoint well, we also need to ensure that the leadership roles to which we appoint have an appropriate focus and workload. I hope we can keep the people pipeline well stocked during my presidency.

Important challenges for the funding of applied mathematics arise from the recent election of Donald Trump in the U.S. and the June 2016 referendum in which the U.K. voted for Brexit. Both events bring about uncertainties regarding the funding of research in their respective countries. SIAM's Committee on Science Policy will have much to keep it busy.

When my term as Vice President at Large ended in 2013, efforts were already underway to prepare for the development of a new SIAM website. Such is the magnitude of the task that the new site is still under development, but it is anticipated that it will be completed in the next few months. SIAM is already making good use of social media, particularly through its Twitter, Facebook, and YouTube accounts. The new *SIAM News* website, launched last year, provides a single source for blog posts, *SIAM News* articles, news stories, videos, etc. about applied math and computational science. Suggestions for articles and blog posts for *SIAM News* are always welcome, and should be sent to managing editor Karthika Swamy Cohen (karthika@siam.org). The new www.siam.org website will complete SIAM's efforts to have a fully up-to-date web presence, and will be much more useful to members, volunteers, and SIAM staff. Stay tuned for more about it.

Nicholas Higham is the Richardson Professor of Applied Mathematics at the University of Manchester. He is the current president of SIAM.

[1] https://www.siam.org/about/board.php

Welcoming Our New Officers

We are pleased to announce the newly-elected members to the SIAM Board and Council, all of whom will serve three-year terms.

Board of Trustees

- Margot Gerritsen, *Stanford University*
- C. Tim Kelley, *North Carolina State University**
- Randall LeVeque, *University of Washington**

Council

- Liliana Borcea, *University of Michigan**
- Per Christian Hansen, *Technical University of Denmark*
- Helen Moore, *Bristol-Myers Squibb*
- Felix Otto, *Max Planck Institute for Mathematics in the Sciences**

* indicates incumbent

siam news

ISSN 1557–9573.

Remembering Sir Christopher Zeeman

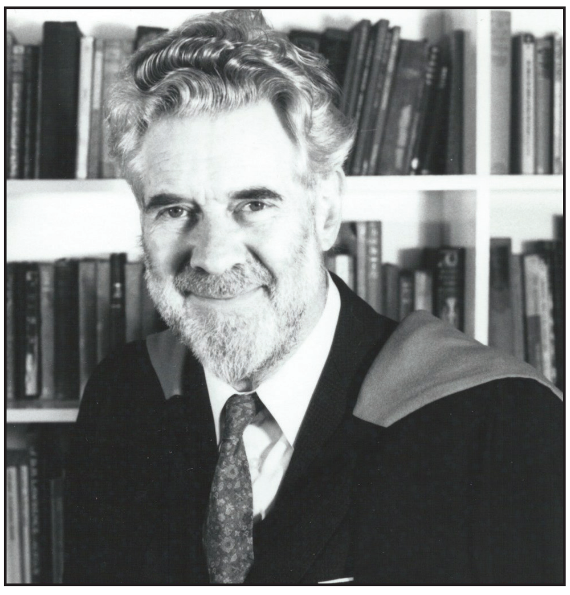
By Tim Poston

Renowned British mathematician Sir Christopher Zeeman passed away in February 2016 at the age of 91. He is best known for his contributions to the fields of geometric topology, singularity theory, and catastrophe theory. One year later, Tim Poston honors his memory.

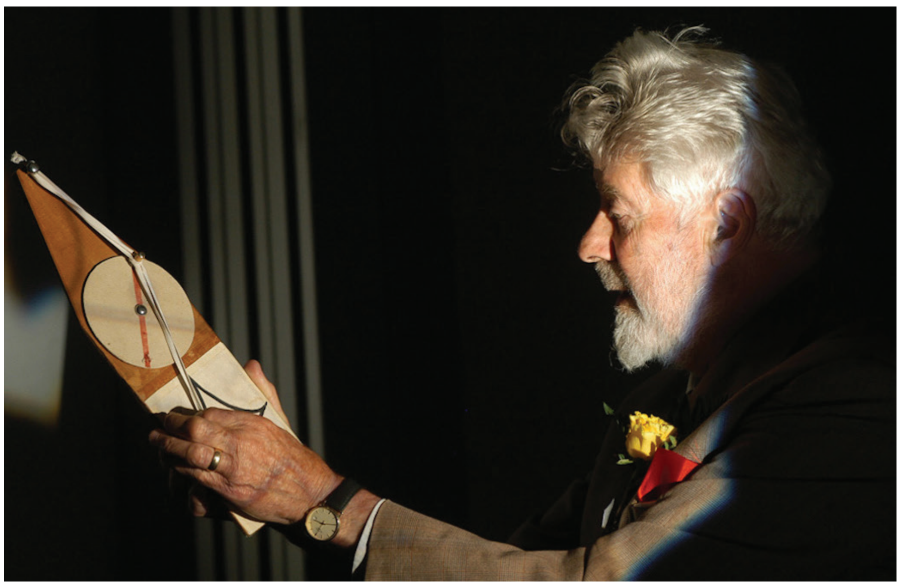
I did not meet Chris Zeeman again after completing my Ph.D. at the University of Warwick under his guidance—I left for Rio, and have wandered widely ever since—but he was always a vivid presence to me. Ian Stewart and I dedicated our book, *Catastrophe Theory and its Applications*, to him, "At whose feet we sit, on whose shoulders we stand;" we still do. Chris was the first faculty member appointed at the University of Warwick. He also founded the Warwick Mathematics Institute in 1965, which remains one of the glories of British mathematics and is now housed in a building that bears his name.

Chris began his research when topology was intensely algebraic, and achieved mastery in 'spectral sequences' (infinite systems of groups developed by Jean Leray in a prisoner-of-war camp) that topology maps into. But Chris's geometric core quickly emerged, with new and deep results on piecewise-linear (PL) topology. To general surprise, this turned out to have essential differences from differential (curved) topology; for example, PL n spheres can always unknot in $(n+3)$ dimensional space, like a curve in dimension 4. Chris also invented 'tolerance spaces'—independently of Henri Poincaré, who had called them 'physical continua'—and applied them to geography and the brain. He was eager to apply 20th century mathematics to all fields, as most 'applied' mathematics at the time occurred in pre-1900 areas.

This desire blossomed when Chris learned René Thom's catastrophe theory, a mixture of deep mathematics and almost metaphysical ideas, from Thom himself. He set out to demystify it for a wide audience and apply it in many fields. The resulting impression that the 'seven elementary catastrophes' were single descriptors caused some controversy in the social sciences; for instance, many of Chris's social and biological models used *one* cusp catastrophe, which his readers often took as a limit. Even with a system truly governed by the bifurcations of a scalar field, where Thom's theorem definitely says the only stably-possible bifurcations for *n* internal variables with (for instance) two-dimensional control are folds and cusps, there is no limit on their number. Indeed, Chris's famous 'catastrophe machine' had four cusps, and there is no reason cellular dynamics should not have four thousand. Thus, his models often showed fruitful and previously-unimagined possibilities for switching behavior, not inevitabilities; however, this was not well understood in the 'soft sciences.' In optics, buckling, laser physics, etc., such confusion did not arise, and the mathematics has continued to yield applications such as efficient simulation.

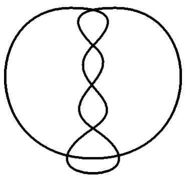


Sir Christopher Zeeman, 1925-2016.



Christopher Zeeman holds a catastrophe machine, a device of his own invention.

Even in n dimensions—certainly in four or five—Chris had an almost tactile feeling for shapes and their changes, and an amazing gift for sharing his findings with both students and colleagues and in his famous public lectures. As his Ph.D. student and teaching assistant, I had the privilege of attending his undergraduate class in topology, which covered a variety of topics. Chris always engaged more intensely than any lecturer I have ever seen, but one day he came in without a greeting, turned his back on the class, and silently began drawing. At the top of the enormous roller blackboard he had previously drawn a row of figures like this one,



identical except for different crossings, drawn as ╳ or ╳. Below the original drawings, Chris reduced each figure step-by-step to simpler forms. For example, if the bottom two crossings showed the small loop to be on one side of the big loop, he pulled it upward, and then untwisted the remaining crossings one by one. Some figures reduced all the way to a circle; some could not. The class watched in silent fascination, until he made a slip while reducing the fourth figure. The whole class loudly objected.

Chris turned to face us. "So you *agree* there's a subject!" he said triumphantly. Nothing could have conveyed more clearly that mathematics is not deduction from arbitrary axioms; axioms serve to capture objects (mental or physical), about which we have real intuitions.

My ears will never again hear him say, "What a lovely geometric argument!" But whenever I find one the phrase echoes in my mind. The news of his death made me realize how much of him, and yet how little, lives on in me and in others he taught.

A profound loss.

Tim Poston has worked in four continents, publishing with co-authors from an archaeologist to a brain surgeon. He has patents from search software to MRI coils, and a recent science fiction novel (with Ian Stewart). He is now chief scientist of a flow simulation startup in Bangalore, India.

Diversity

Continued from page 1

the symmetric state. Both requirements are consistent with, but do not follow from, Curie's principle. With regard to the first requirement, it is important to note that Curie's principle pertains to exact symmetries; it says nothing about cases involving approximate symmetries and, hence, about the stability of the states. Indeed, it is not true that nearly symmetric causes (which are also determined by the initial conditions) will generally lead to nearly symmetric effects, as already demonstrated by the phenomenon of spontaneous symmetry breaking. This is the very reason why the symmetric state is not realized in the scenarios considered here, despite the symmetry of the system. Concerning the second requirement, while it might be counterintuitive that the system should be asymmetric in order for the symmetric state to be stable, Curie's principle provides no *a priori* reason why an asymmetric cause could not produce a symmetric effect. Note that in AIS it is not the existence of a stable symmetric state for an asymmetric system that is striking, but instead the fact that such a state can only be stable when the system is asymmetric.

The relation between AIS and symmetry does not end there. In AIS, it is the individual realization of the system whose symmetry must be broken to preserve the symmetry of the solution. The symmetry of the solution is reflected, however, in the region of the parameter space defined by the ensemble of all possible systems for which the symmetric solution is stable (see Figure 2). This too is the converse of what is observed for symmetry breaking, where the realized stable solution does not have the symmetry of the system, but the set of all stable solutions does.

Symmetry, as a mathematical concept, has foundational implications in many fields [3]. In physics, Hermann Weyl noted over 60 years ago that all *a priori* statements have their origin in symmetry [4]; if Frank Wilczek's predictions are of any guidance, this trend will not change over the next 100 years [5]. In particular, symmetry breaking is expected to continue playing a significant role in allowing symmetric theories to explain asymmetric observations. Complementarily, the phenomenon of AIS shows that asymmetric theories and models may be required to describe symmetric realities, which by itself ought to raise questions about assumptions often tacitly made on the causes when the effects are symmetric.

Lastly, the notion that symmetry (or lack thereof) can lead to surprising collective behavior is also appealing to non-researchers, and this has been explored in an outreach project with high school students in a dance piece titled *Syncing Up Without Sameness*[1] (see Figure 3).

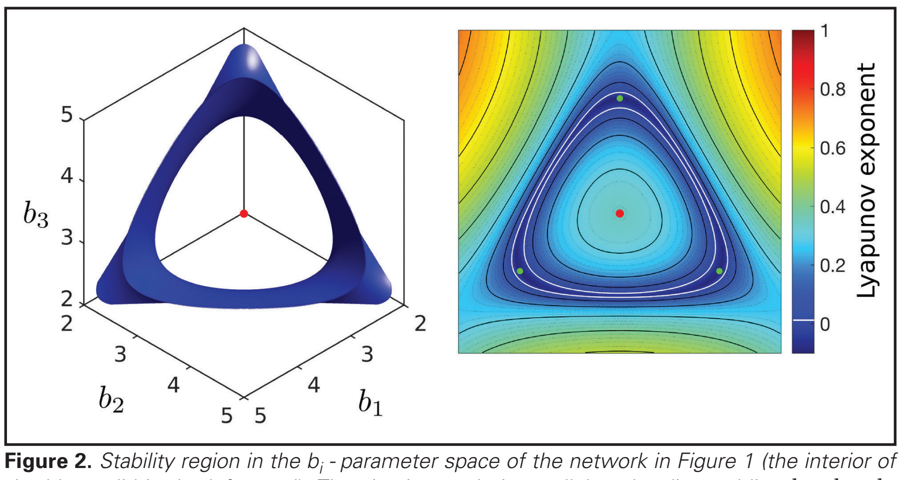


Figure 2. *Stability region in the b_i - parameter space of the network in Figure 1 (the interior of the blue solid in the left panel). The viewing angle is parallel to the diagonal line $b_1 = b_2 = b_3$ (indicated by the red dot). The right panel shows the Lyapunov exponent on the plane that is perpendicular to the diagonal line and contains all global minima (green dots) of the Lyapunov exponent. Image credit: Takashi Nishikawa and Adilson E. Motter.*

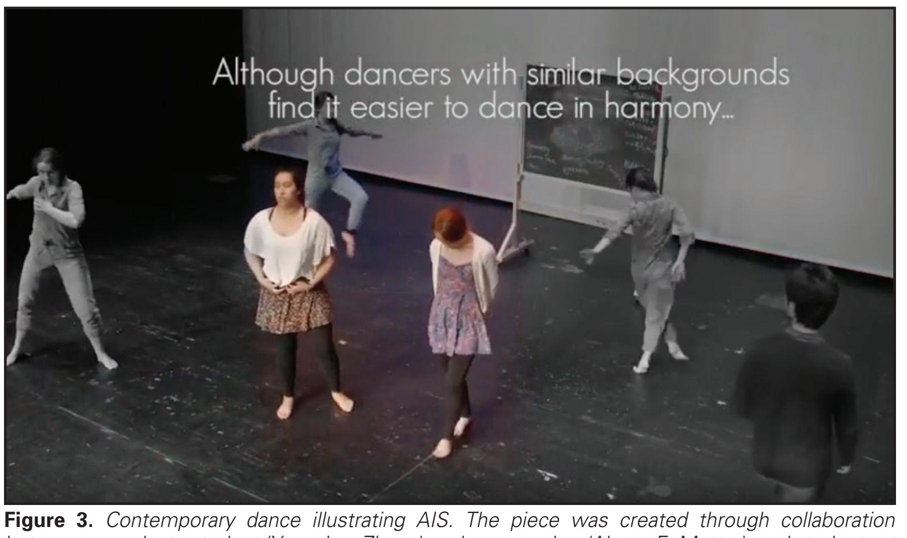


Figure 3. *Contemporary dance illustrating AIS. The piece was created through collaboration between a graduate student (Yuanzhao Zhang), a choreographer (Alyssa E. Motter), and students at Regina Dominican High School. Credit: Yuanzhao Zhang, Alyssa E. Motter, and Adilson E. Motter.*

[1] https://www.youtube.com/watch?v=qFCX2pZXDBg

References

[1] Curie, P. (1894). Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique. *J. Phys. Theor. Appl., 3*, 393.

[2] Nishikawa, T., & Motter, A.E. (2016). Symmetric states requiring system asymmetry. *Phys. Rev. Lett., 117*, 114101.

[3] Stewart, I., & Golubitsky, M. (1992). *Fearful Symmetry: Is God a Geometer?* Cambridge, MA: Blackwell Publishers.

[4] Weyl, H. (1952). *Symmetry.* Princeton, NJ: Princeton University Press.

[5] Wilczek, F. (2016). Physics in 100 years. *Physics Today, 69*, 32.

Takashi Nishikawa is a research associate professor and Adilson Motter is the Charles E. and Emma H. Morrison Professor of Physics and Astronomy at Northwestern University.

SIAM and the National Science Foundation

By Jim Crowley

SIAM is a member of many umbrella organizations. Along with its sister societies, SIAM participates in the Joint Policy Board for Mathematics (JPBM), the Conference Board in the Mathematical Sciences (CBMS), the Computing Research Association (CRA), and many others.

At its October meeting in Washington, D.C., the JPBM heard from Deborah Lockhart, Deputy Assistant Director of the Directorate for Mathematical and Physical Sciences (MPS) at the National Science Foundation (NSF). She offered noteworthy suggestions for SIAM and the larger mathematical sciences community.

Lockhart noted that as Michael Vogelius nears the end of his term as head of the Division of Mathematical Sciences (DMS) in January 2018, the SIAM community can help identify new candidates for this important position. The DMS could also use help with its search for program directors, as half of the scientific staff tend to be rotators who serve for limited periods of time. The NSF needs the community's input to find suitable individuals for these roles.

The NSF could also benefit from the scientific community's assistance to identify and spread the word about important, new, and developing scientific areas that require research funding. In addition, the organization is seeking suitable individuals to serve on advisory boards, like the MPS Advisory Committee, and to organize and attend workshops. These workshops are meant to develop ideas and define and establish NSF-wide initiatives and priorities that show how the mathematical sciences can contribute.

Furthermore, opportunities exist for research funding in cross-cutting areas where the mathematical sciences—and members of the SIAM community—can make important contributions. Examples of these areas include cybersecurity and big data, and solicitations such as Designing Materials to Revolutionize and Engineer our Future (DMREF) and Innovations at the Nexus of Food, Energy and Water Systems (INFEWS). Lockhart also encouraged increased participation from mathematicians in the NSF Graduate Fellowship program.[1] The mathematical sciences have traditionally received fewer of these fellowships than one might think, due to a relatively low proportion of submitted applications. The algorithm has changed somewhat, but the mathematical sciences are still underrepresented – more applicants are needed from the discipline.

Moreover, Lockhart encouraged participation in programs such as Inclusion across the Nation of Communities of Learners of Underrepresented Discoverers in Engineering and Science (INCLUDES),[2] Increasing the Participation and Advancement of Women in Academic Science and Engineering Careers (ADVANCE)[3] and NSF Research Traineeship (NRT).[4]

The NRT program is a multidisciplinary effort designed to facilitate the development and implementation of new and potentially-transformative models for science, technology, engineering, and mathematics (STEM) graduate education training. Its goal is to ensure that graduate students develop the skills, knowledge, and competencies needed to pursue a range of STEM careers. NRT encourages collaboration across disciplines, and the program's emphasis changes year to year; data science was a recent emphasis.

Lockhart urged the community to hold workshops on new and emerging areas as well as new federal priorities to help define emergent topics and demonstrate the mathematical sciences' role in these disciplines. Since solicitations are announced quickly with short turnaround times, advance discussions are important.

In short, Lockhart made an eloquent plea to help the NSF – a message that could be replicated across all agencies that fund scientific research relevant to the SIAM community.

1 https://www.nsfgrfp.org/
2 https://www.nsf.gov/news/special_reports/nsfincludes/index.jsp
3 https://www.nsf.gov/crssprgm/advance/
4 https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505015

Jim Crowley is the executive director of SIAM.

Data-driven Discovery

Continued from page 1

the elliptical nature of planetary orbits (see Figure 1, on page 1). Newton built upon this work, proposing the existence of gravity in order to derive $\mathbf{F}=m\mathbf{a}$ and explain Kepler's elliptic orbits. A cynic could argue that Newton provided nothing new in terms of prediction, just a more convoluted way to derive elliptical orbits through calculus. However, Newton used the inferred $\mathbf{F}=m\mathbf{a}$ to facilitate the development and characterization of new systems never before considered or observed. The discovery of this fundamental governing law was critical for technological development and enabled unprecedented engineering and scientific progress, such as sending a rocket to the moon.

The success of Newton's calculus fueled the scientific revolution and led to many of the canonical models of mathematical physics, including the heat equation, wave equation, Poisson's equation, and Navier-Stokes equations, among other significant developments. But for many modern applications, governing equations are often unknown or only partially known, and exhibit strong nonlinearities, parametric dependencies, multi-scale phenomena, intermittency, and/or transient behavior. Such systems often take the general mathematical form

$$u_t = N(u,x,t;\mu). \qquad (1)$$

The function $N(\cdot)$ is an unknown right-hand side that describes the ordinary or partial differential equation in terms of u, its derivatives, and parameters in μ. Our objective is to *discover* $N(\cdot)$ given only time-series measurements of the system (see Figure 2). A key assumption is that the true $N(\cdot)$ is comprised of only a few terms, making the model sparse in the space of all possible combinations of functions. For example, Burgers' equation ($N=-uu_x+\mu u_{xx}$) and the harmonic oscillator ($N=-i\mu x^2u-i\hbar u_{xx}/2$) each have only two terms. This is consistent with Occam's razor: the most likely governing equation is the simplest one that works.

Naïve approaches to discovering $N(\cdot)$ lead to a combinatorially large search through all possible models. To overcome this difficulty, researchers have developed methods spanning ideas from nonlinear regression to artificial neural networks [4, 6, 14, 17]. More recently, Michael Schmidt and Hod Lipson used a genetic algorithm to *distill* the free-form laws from measurements in a seminal contribution [13]. It is important to note that the most parsimonious model was selected by balancing model complexity and accuracy via Pareto analysis [10]. An alternative approach uses emerging sparse regression techniques to determine $N(\cdot)$ without an intractable (*np*-hard) combinatorial brute-force search [2]. Specifically, a library $\mathbf{\Theta}(\mathbf{U})$ of candidate linear, nonlinear, and partial derivative terms for the right-hand side is constructed. Each column of $\mathbf{\Theta}(\mathbf{U})$ contains the values of a candidate term evaluated using the collected data. In this library, one can write the dynamics as

$$\mathbf{U}_t = \mathbf{\Theta}(\mathbf{U})\xi, \qquad (2)$$

where $\mathbf{U}_t$ is a vector of time derivatives of the measurement data and ξ is a sparse vector, with each nonzero entry corresponding to a functional term to be included in the dynamics. Finding the sparsest vector ξ consistent with the measurement data is now feasible with advanced methods in sparse regression, which makes it possible to find the most parsimonious model while circumventing the combinatorial search. Moreover, this approach has found success in a wide variety of ordinary [2, 9] and partial differential equation [11] settings. One can collect the time-series measurements from either an Eulerian framework where the sensors are fixed spatially, or in a Lagrangian framework where the sensors move with the dynamics [11]. This method is part of a growing effort to leverage sparsity in dynamical systems [3, 12, 15].

The initial success of these methodologies, including sparse regression and genetic algorithms, suggest that one can integrate many concepts from statistical learning with traditional scientific computing and dynamical systems theory to discover dynamical models from data. This integration of nonlinear dynamics and machine learning opens the door for principled versus heuristic methods for model construction, nonlinear control strategies, and sensor placement techniques. Additionally, these new model identification methods have transformative potential for parameterized systems and multiscale models where first principle derivations have remained intractable, such as neuroscience, epidemiology, and the electrical grid.

References

[1] Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science, 16*(3), 199-231.

[2] Brunton, S.L., Proctor, J.L., & Kutz, J.N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences, 113*(15), 3932-3937.

[3] Brunton, S.L., Tu, J.H., Bright, I., & Kutz, J.N. (2014). Compressive sensing and low-rank libraries for classification of bifurcation regimes in nonlinear dynamical systems. *SIAM Journal on Applied Dynamical Systems, 13*(4), 1716-1732.

[4] Crutchfield, J.P., & McNamara, B.S. (1987). Equations of motion from a data series. *Complex Systems, 1*, 417-452.

[5] Donoho, D.L. (2015). 50 years of data science. In *Tukey Centennial Workshop*. Princeton, NJ: Princeton University.

[6] Gonzalez-Garcia, R., Rico-Martinez, R., & Kevrekidis, I.G. (1998). Identification of distributed parameter systems: A neural net based approach. *Computers & Chemical Engineering, 22*, S965-S968.

[7] Hey, T., Tansley, S., & Tolle, K.M. (Eds.) (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. (Vol. 1). Redmond, WA: Microsoft Corporation.

[8] Kepler, J. (1627). *Tabulae Rudolphinae, quibus astronomicae scientiae, temporum longinquitate collapsae restauratio continetur*. Ulm, Germany: Vlmae, Typis J. Saurii.

[9] Mangan, N.M., Brunton, S.L., Proctor, J.L., & Kutz, J.N. (2016). Inferring biological networks by sparse identification of nonlinear dynamics. arXiv:1605.08368. To appear in *IEEE Transactions on Molecular, Biological, and Multi-Scale Communications.*

[10] Pareto, V. (1964). *Cours d'e´conomie politique* (Vol. 1). Geneva, Switzerland: Librairie Droz.

[11] Rudy, S., Brunton, S.L., Proctor, J.L., & Kutz, J.N. (2016). Data-driven discovery of partial differential equations. arXiv:1609.06401.

[12] Schaeffer, H., Caflisch, R., Hauck, C.D., & Osher, S. (2013). Sparse dynamics for partial differential equations. *Proceedings of the National Academy of Sciences, 110*(17), 6634-6639.

[13] Schmidt, M., & Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science, 324*(5923), 81-85.

[14] Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science, 338*(6106), 496-500.

[15] Tran, G., & Ward, R. (2016). Exact recovery of chaotic systems from highly corrupted data. arXiv:1607.01067.

[16] Tukey, J.W. (1962). The future of data analysis. *The Annals of Mathematical Statistics, 33*(1), 1-67.

[17] Voss, H.U., Kolodner, P., Abel, M., & Kurths, J. (1999). Amplitude equations from spatiotemporal binary-fluid convection data. *Physical review letters, 83*(17), 3422.

Steven L. Brunton is an assistant professor of mechanical engineering, adjunct assistant professor of applied mathematics, and a data science fellow with the eScience Institute at the University of Washington. J. Nathan Kutz is professor of applied mathematics, adjunct professor of physics and electrical engineering, and a senior data science fellow with the eScience Institute at the University of Washington. Joshua L. Proctor is an associate principal investigator with the Institute for Disease Modeling as well as affiliate assistant professor of applied mathematics and mechanical engineering at the University of Washington.

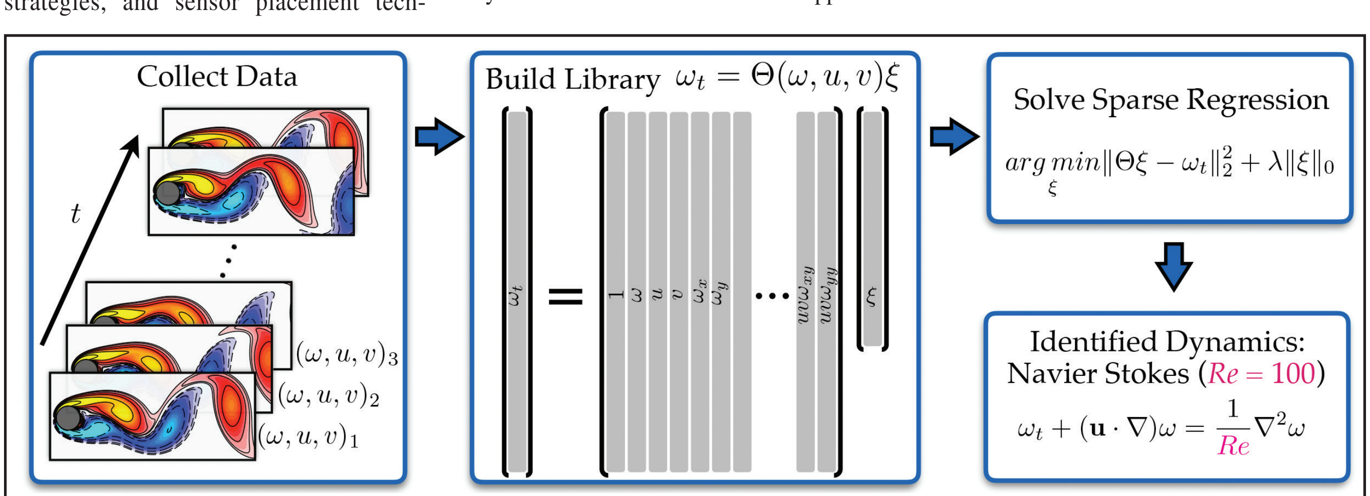


Figure 2. *Data-driven discovery algorithm in which time-series data alone is used to construct the governing equations of the measured system. Image credit: Samuel Rudy, Steven L. Brunton, J. Nathan Kutz, and Joshua L. Proctor.*

Systems Analysis to Inform and Support Global Transformations

By Stephen M. Robinson, Elena Rovenskaya, and Ulf Dieckmann

Governments and private decision-makers worldwide now confront problems of unprecedented difficulty. The challenges include the increasing scale and coupling of complex systems, the acceleration of technological advances, economic interactions, and information flows.

New Kinds of Global Challenges

Four relatively new trends heighten the difficulty of the aforementioned problems. First, the *increasing scale* of the world's population and people's activities—and hence their impact on the natural environment—runs the risk of exceeding planetary boundaries. Secondly, *interdependencies* among people, companies, and countries have grown to the extent that local failures in public services can create mass emergencies; the systemic risk underlying the latest global financial crisis is a prominent example. Thirdly, the high speed of *technological advances* presents challenges to long-term planning, such as the planning of investments in large infrastructures subject to high uncertainties. And lastly, the advent of new social media facilitates mass agitation over contentious issues, often leading to *irrational politicization*.

Traditional technical education in operations research (OR) does not provide sufficient tools for assisting decision-makers in handling these problems. It is still common to analyze systems via a quantitative model for predicting their future—either deterministically or stochastically—and then examine the ways in which a single, known criterion for the goodness of a solution varies with decisions. Students in traditional university programs learn effective technical methods to make such analyses, producing what are often called "technocratic" solutions. These solutions are important and useful, but insufficient to serve as guidelines for handling problems of the kinds described above.

For one thing, decentralized decision-making under bounded rationality is a characteristic of many of these problems. One must account for this, as well as for the aforesaid interdependencies, to produce feasible solutions. For instance, enhancing transportation to improve an area's economic condition will not work if those influential in the local government block the new arrangement to preserve their monopoly on transportation. Likewise—an actual, recent example—a program that builds wells to provide clean drinking water in rural areas will fail if villagers are not both able and motivated to keep the pumps in good repair.

A typical OR graduate is unlikely to see the real problem in such situations, because the key dimensions in these examples are not primarily technical, and often not even economic, but rather social and/or political. Moreover, even if the graduate is able to clearly recognize the additional dimensions, he/she will not typically possess a toolkit of skills for tackling them. In addition, the graduate will often not know how to carry on a constructive conversation with stakeholders in terms they can understand.

Systems Analysis: A Multidimensional Tool

Over the past decades, the field of systems analysis has matured into a broadly applicable tool for the development of integrated multidisciplinary solutions and policy advice for some of the world's most pressing problems. From its inception in military analyses during World War II to its extension to civic applications pioneered by the RAND Corporation, modern systems analysis employs concepts, models, and methods that help account—simultaneously and as seamlessly as possible—for a problem's interwoven technical, economic, social, political, and communication dimensions.

Systems analysis looks across borders and sectors to identify feedbacks, trade-offs, and synergies. For this purpose, it builds on and interlinks sectoral and disciplinary approaches to enable holistic and global outlooks. In doing so, it strives to recognize uncertainties, nonlinearities, adaptive capacities, tipping points, bounded rationality, and normative pluralism. Systems analysis helps identify smart pathways through the complex nexus of interdependent processes to reach a world that accommodates the needs and aspirations of different groups and respects the limits imposed by the planet itself.

CAREERS IN MATHEMATICAL SCIENCES

For the last 45 years, the International Institute for Applied Systems Analysis (IIASA) in Laxenburg, Austria, has led developments in systems analysis, continually enhancing the field's methodologies and strengthening its applications. Problems of global and universal relevance lie at the heart of IIASA's research agenda, including sustainable development, climate change, energy strategies, environmental protection, resource utilization, land-use change, ecosystem management, risk and resilience, and population growth. Recent impacts of IIASA's research range from shaping European Union air pollution policy to providing results central to the Paris climate summit negotiations.

New Challenges Require New Preparation

Dealing with the aforementioned global and universal problems requires access to skills in multiple dimensions covered by systems analysis:

- **Technical skills:** Our current educational sector does this quite well.
- **Economic skills:** How to assess costs and benefits, both direct and indirect?
- **Sociological skills:** What solutions are acceptable in a given situation? Why and how can such acceptability be changed?
- **Political skills:** Who will stand in the way of certain solutions? Why and how can such obstacles be overcome?
- **Communication skills:** How can we talk with stakeholders to promote mutual understanding?

Training graduates across this skill profile may seem hopelessly unrealistic. OR students already spend much time learning about the technical dimension in their study area, which is why they become good at it. There is no way to formally teach them—at a similar depth and breadth—economics, sociology, politics, and communication.

The key is to recognize that "access to" does not necessarily mean "mastery of." What we should be able to do is teach the students both how to work with experts in these other dimensions, and why such collaboration is necessary to develop solutions to problems arising in complex systems. As yet, few OR programs do this effectively. Some do not even realize that the problem exists.

Focus-group discussion with local farmers who are non-beneficiaries from the considered irrigation site in Ethiopia. © Christoph Perger (IIASA).

A Practical Way Ahead

How might we change this situation? A helpful role model might be the "capstone" courses taken shortly before graduation by students in many engineering programs. In these courses, students work in teams to solve real-world problems in their respective disciplines, often in cooperation with industries or government agencies. This helps prepare them for the kind of situations they will face after graduation.

If we want to build competence in teamwork when not all team members are from technical disciplines, it makes sense to do so by having students work to solve a complex, multidimensional real-world problem under the guidance of experts skilled in multidisciplinary research. In fact, this is quite feasible; the Young Scientists Summer Program (YSSP) at IIASA has utilized this approach for nearly 40 years, bringing together about 50 young scientists for three months each summer to work in this format as a multidisciplinary group.

The YSSP has been very successful, but it is small and cannot possibly provide as many skilled systems analysts as will be required for dealing with current global and universal problems. However, a revision and expansion of capstone courses in strong OR departments all over the world, to train students for work in multidisciplinary teams on complex problems, could significantly enlarge that pool of analysts. There is surely no lack of such problems for them to solve.

Where can we find the students graduating in other disciplines to complete the needed teams? Here it is helpful to think about the career challenges graduates in sociology, political science, communication, or other disciplines currently face. The job markets in those areas are often much less promising than those in the science, technology, engineering, and mathematics (STEM) fields. Would not some of those graduates find interest and appeal in contributing to solutions for some of the most difficult problems facing humanity? And would not an effort like this help build bridges among very different disciplines, leading to new perspectives on those problems – perspectives that we would never see if we do not leave our silos?

Universities and their schools are very unlikely to take the initiative for a change of this magnitude, but professional societies are in a good position to lead. Through meetings and publications they set the standard for what is currently important, thereby putting pressure on academic programs that otherwise might not recognize the need to change. By thus transforming the conversation, professional societies could play a critical role in advancing competence in techniques that humanity already needs now, and will need even more in the future.

Portions of this article appeared in the "President's Desk" column in the December 2014 issue of OR/MS Today. They are reused with permission.

Stephen M. Robinson is professor emeritus in the Department of Industrial and Systems Engineering at the University of Wisconsin-Madison, and a research scholar of the International Institute for Applied Systems Analysis (IIASA). He is an elected member of the National Academy of Engineering, a fellow of SIAM, and a fellow and former president of the Institute for Operations Research and the Management Sciences (INFORMS). Elena Rovenskaya is the director of the Advanced Systems Analysis (ASA) Program at IIASA and a research scholar at the Department of Optimal Control of the Faculty of Computational Mathematics and Cybernetics at Lomonosov Moscow State University in Russia. Her scientific interests include theory and applications of optimal control, as well as economic-environmental modeling. Ulf Dieckmann is director of the Evolution and Ecology Program (EEP) at IIASA. He is working on the theory of adaptive dynamics, game theory, network theory, speciation theory, life-history theory, evolutionary ecology, spatial ecology, social evolution, and fisheries-induced evolution.

In cooperation with the International Fund for Agricultural Development (IFAD), the International Institute for Applied Systems Analysis (IIASA) supports the impact evaluation of IFAD-funded projects by collecting data from farmers in the field, to be fed into models supporting the development of future policy scenarios. The image shows an IFAD-funded irrigation site in Ethiopia. © Christoph Perger (IIASA).

What are the Odds?

Looking at Coincidences Mathematically

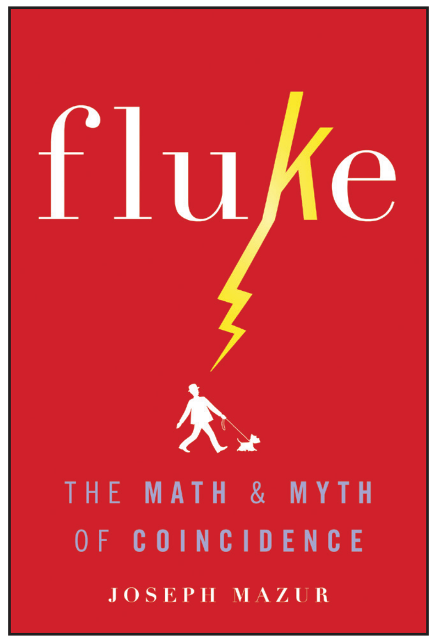
Fluke: The Math and Myth of Coincidence. *By Joseph Mazur. Basic Books, New York, NY, March 2016. 288 pages, $26.99.*

BOOK REVIEW
By James Case

The subject of Joseph Mazur's fifth popular math book is "seemingly unlikely events." He recognizes three main types—coincidences, flukes, and serendipities—and offers relatively concise definitions for each. Sandwiched between a brief "Introduction" and an equally-brief "Epilogue" are four main parts: "The Stories," "The Mathematics," "The Analysis," and "The Head Scratchers."

Part I, consisting of three chapters, introduces the ten "real-life stories" on which the greater part of the book is based. They range in complexity from that of a Texas woman who won four multimillion dollar lotteries in less than twenty years to the tale of two men named Francesco who arrived in the same hotel lobby at the same time, expecting to meet and interview a stranger named Manuela for possible employment. Both met a woman of that name, retired to a nearby conference room, and began a somewhat bewildering conversation before returning to the lobby and discovering that each Francesco had been talking to the wrong Manuela!

Part II contains five chapters on the computation of probabilities of compound events. Mazur offers many examples, including the following: diffusion processes, laws of large numbers, Bernoulli trials, the birthday problem, Pascal's triangle, Galton's peg board simulation of an unbiased one-dimensional random walk, Huxley's typewriting monkeys, and the proverbial flap of a butterfly's wing. He addresses all this and more without invoking anything beyond grade-school arithmetic and elementary algebra.



Fluke: The Math and Myth of Coincidence. By Joseph Mazur. Courtesy of Basic Books.

Part III, which contains only two chapters, applies the lessons learned in Part II to the analysis of the stories from Part I to demonstrate that, whereas some seemingly unlikely events are indeed highly unlikely, others should not even be considered surprising. In particular, given the large number of high-value lotteries in the world and the vast number of people who participate religiously, it is all but inevitable that someone somewhere will win more than once, and by no means unlikely that there will be a four-time winner!

Part IV of *Fluke* is perhaps the most interesting. It consists of five essays concerning coincidences that, at least for the moment, completely escape analysis. The first investigates coincidences in DNA evidence gathered at crime scenes and lawyerly attempts to mislead jurors regarding the likelihood of consequent mistakes. The second describes accidental findings by scientists studying seemingly unrelated phenomena, such as the discovery of penicillin in Alexander Fleming's untidy laboratory, where fungus from a separate study inadvertently contaminated a staphylococcus culture, and Röntgen's discovery of X-rays while investigating electrical currents in a partially evacuated glass tube. According to Mazur, any number of earlier investigators also produced such rays, but failed to notice them due to weaker currents and/or less complete vacuums. The third essay involves a rogue trader who wagered recklessly while failing to anticipate two massive flukes. The first one made him rich, while the second left him bankrupt. The fourth considers attempts to evaluate psychic powers and extrasensory perception (ESP) statistically, while the fifth essay compares some elaborate coincidences in literary fiction with their (typically less-elaborate) counterparts in real life.

Books of the present sort can serve as texts for the "quantitative reasoning" courses colleges and universities are increasingly obliged to offer victims of "math anxiety." Intended to enhance a student's ability to assess the risks and uncertainties encountered in daily life, such courses can prove immediately useful to novices faced for the first time with the need to purchase healthcare insurance, schedule a holiday picnic, or acquire additional student loans.

James Case writes from Baltimore, Maryland.

Sci-Hub: Stealing Intellectual Property or Ensuring Fairer Access?

By Ted Lockhart

For anyone who wants to read or download an article in a scientific journal, Sci-Hub is one-stop shopping. During the past five years, thousands of scientists, students, and others have used the internet search engine and repository of scientific journal articles, which currently stores approximately 50 million articles in its online database and is capable of locating and accessing virtually any article that exists online.

Sci-Hub was the idea and creation of Alexandra Elbakyan, a university student from Kazakhstan. As a student, she found that her university's library did not carry many of the journals she needed for her research. When she tried to access them online, she often ran up against paywalls charging fees of $30 or more for a single article. For a student from a relatively poor country, this was out of the question [4]. Hence, Elbakyan decided to design a search engine that would employ a large collection of usernames and passwords. Collectively, they would enable users to access a vast array of online scientific journals while bypassing unexpected, unwanted paywalls. Elbakyan called her creation "Sci-Hub."

How does Sci-Hub acquire the usernames and passwords that it stores? Many are provided voluntarily by individuals who are sympathetic to the Sci-Hub cause. Elbakyan says little about Sci-Hub's sources except to deny that she or Sci-Hub's maintainers obtain them by phishing, i.e., using subterfuge to steal them from their owners. However, she admits that some of the volunteered usernames and passwords may have originally been phished by their providers [6].

Understandably, commercial publishers strongly oppose Sci-Hub. The largest of them, Elsevier, obtained a temporary injunction in Federal District Court in October 2015 ordering Sci-Hub and Elbakyan to cease distribution of any of Elsevier's copyrighted works [3]. The stakes were extraordinarily high for the defendants, since their potential liability ranged from $750 to $150,000 *for each pirated journal article*. Elbakyan chose not to take her chances in court and is now in seclusion, rumored to be somewhere in Russia. Meanwhile, Sci-Hub continues to operate despite publishers' strenuous efforts to disable it [5].

Many professional scientific and engineering societies also strongly oppose Sci-Hub. These organizations rely on proceeds from their proprietary journals to help finance their activities. Sci-Hub enables individuals to circumvent paywalls that restrict access to these journals and thus deprives professional societies of a vital source of revenue.

Publishers argue that most people who use Sci-Hub, especially in wealthier countries, do so as a matter of convenience rather than an inability to afford access fees. They call attention to the special discounts they offer some of their customers, and note that scientists can get free copies of articles directly from their authors or through interlibrary loan. However, their critics contend that most consumers do not qualify for publishers' special deals. Furthermore, contacting authors directly has serious drawbacks, as in the case of a student from Iran:

> Just as spring arrived last month in Iran, Meysam Rahimi sat down at his university computer and immediately ran into a problem: how to get the scientific papers he needed. He had to write up a research proposal for his engineering Ph.D. at Amirkabir University of Technology in Tehran . . .
>
> But every time he found the abstract of a relevant paper, he hit a paywall . . . He looked at his list of abstracts and did the math . . . Purchasing the papers was going to cost $1000 this week alone—about as much as his monthly living expenses—and he would probably need to read research papers at this rate for years to come . . .
>
> Many academic publishers offer programs to help researchers in poor countries access papers, but only one . . . seemed relevant to the papers that Rahimi sought. It would require him to contact authors individually to get links to their work, and such links go dead 50 days after a paper's publication. The choice seemed clear: Either quit the Ph.D. or illegally obtain copies of the papers [1].

Elbakyan argues that commercial academic publishers exploit students, scientists, and others. Publishers receive excessive profit margins, as high as 35%, by using very aggressive business practices. Academic publishers do not generate the scientific results reported in their journals, yet they reap exorbitant financial returns. Because scientists' career success often depends on having their articles published in certain "high impact" journals, they have no choice but to accept publishers' terms. In Elbakyan's view, the current regime is both unfair and ineffective and should be replaced by an open-access model.

What about the ethics of Sci-Hub? Elbakyan has been accused of thinking that she is above the law [2]. Her accusers say that Sci-Hub facilitates the theft of publishers' intellectual property, and stealing is unethical – case closed. For her part, Elbakyan denies the legitimacy of laws that regard scientific knowledge as private property. She cites Article 27 of the *Universal Declaration of Human Rights*: "Everyone has the right freely to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits." However, whether so general a principle can justify so specific an enterprise as Sci-Hub is debatable.

Elbakyan's main complaint against Elsevier and other commercial publishers is that their paywalls and exorbitant access fees exploit economically-disadvantaged individuals and organizations. This strong sense of injustice is often compounded by the feelings of discouragement that aspiring scientists experience because of obstacles to assessing current research in their fields. Elbakyan believes that Sci-Hub has helped bring this sad state of affairs to the public's attention and given hope that a fairer, more fruitful open-access regime will eventually emerge. Of course, open-access is not a panacea and raises its own set of questions, including whether and how to use peer-review.

Academic publishers feel that they are simply playing by the rules of the established system and have gone beyond the call of duty in many ways. However, it is the morality of that system and of Sci-Hub's revolt against it that is the issue here. It should not surprise us that many scientists who use Sci-Hub feel no shame in doing so. It is difficult not to sympathize at least a little with Elbakyan and her fellow "pirates" who created and use Sci-Hub to do the scientific research that society delegates to them. Perhaps Sci-Hub will turn out to be the stimulus that inspires a fairer, more effective system of access to scientific knowledge, one that is good for both scientists themselves and those for whom science holds vital importance.

References

[1] Bohannon, J. (2016, April 28). Who's downloading pirated papers? Everyone. *Science*, *352*(6285), 508-512. doi:10.1126/science.aaf5664

[2] Cochran, A. (2016, February 25). A Funny Thing Happened on the Way to OA [Web blog post]. *The Scholarly Kitchen*. Retrieved from https://scholarlykitchen.sspnet.org/2016/02/25/a-funny-thing-happened-on-the-way-to-oa/

[3] Elsevier Inc. et al v. Sci-Hub et al, 1:15-cv-04282-RWS (New York Southern District Court 2016).

[4] Rosenwald, M.S. (2016, March 30). This student put 50 million stolen research articles online. And they're free. *Washington Post*. Retrieved from https://www.washingtonpost.com/local/this-student-put-50-million-stolen-research-articles-online-and-theyre-free/2016/03/30/7714ffb4-eaf7-11e5-b0fd-073d5930a7b7_story.html?utm_term=.aae1484b1fb7

[5] Schiermeier, Q. (2015, December 4). Pirate research-paper sites play hide-and-seek with publishers. *Nature*. doi:10.1038/nature.2015.18876

[6] Taylor, M. (2016, February 25). Does Sci-Hub phish for credentials? [Web blog post]. *Sauropod Vertebra Picture of the Week*. Retrieved from https://svpow.com/2016/02/25/does-sci-hub-phish-for-credentials/

Ted Lockhart is emeritus professor of philosophy at Michigan Technological University. He is the author of Moral Uncertainty and Its Consequences (Oxford Univ. Press, 2000).

Supporting Diversity in Mathematics Departments

By Rosalie Bélanger-Rioux

Most mathematics departments are not very diverse. Many mathematical sciences and mathematics education faculty would like this to change, but feel lost as to what they can do as individuals. This challenge motivated me to organize "Increasing Diversity and Inclusion in Mathematics: Some Inspiring Initiatives,"[1] a minisymposium at the inaugural SIAM Conference on Applied Mathematics Education (ED16) last September.

As part of the minisymposium, I described various initiatives I started with colleagues from Harvard University's Mathematics Department, demonstrating that individual projects can collectively make a substantial difference. Among the various ideas presented for individuals to pursue were a welcome lunch for new math majors, a newsletter, informational events about possible courses or summer research, and sharing a collection of experiences of women in the department to inspire other women. Some of these propositions are specifically aimed at students who may lack a strong social support network and access to information in their math departments. There is strength in numbers and various ways to find allies, be they faculty or students, in pursuing such projects. For example, some of our faculty attended a "Gender Gap in Math" event organized by students, which helped them show support and develop ties with the students. This event also facilitated collaborations on various initiatives. Additionally, departments can organize diversity training for faculty and graduate students, and use it to advertise a diversity discussion group. Such programs grow as more people engage with the issue and offer each other support.

A training led by Cynthia Anhalt (University of Arizona) and myself encouraged participants to read case studies of underrepresented students. The business and teaching communities often use case studies to reflect on real situations and prepare businesspeople or teachers to respond to their employees, students, etc. Participants discussed details of the cases in breakout groups before sharing their thoughts with all attendees. For example, one case described the experience of Luis, a first-generation Latino student going back to school as a mathematics major while working part-time as a restaurant manager to support his family. Luis is on track to obtain a C in his course, mostly because of homework. He does not know, and cannot relate much to, other students in the class, has no one to collaborate with on assignments, and feels that others do not take his contributions in class seriously because they see him as a "C student." He plans to drop out of school. Training participants were asked to reflect on what they could do as faculty members to support Luis and make him feel like he belongs.

Knowing how to act in the face of delicate interactions, especially related to issues of diversity, is often challenging. I imagine many faculty would like to say something positive in such cases, but are challenged by the prospect of coming up with a suitable response in the moment. Some reactions might be appropriate, supportive, or welcoming, while others might unwittingly be alienating or hurtful. It is crucial to maintain a safe and respectful place where we as educators understand that this is a common problem; many of us grapple with these issues, may unintentionally say things we regret, and ultimately learn from the experiences. Such an understanding allows us to begin (or continue) to reflect on our thoughts and actions, our biases and fears, and our best intentions for students and colleagues.

From left to right: Edward Doolittle, Rosalie Bélanger-Rioux, Cynthia Anhalt, and Rachel Levy at the 2016 SIAM Conference on Applied Mathematics Education. Photo credit: Rachel Levy.

Developing culturally-relevant curriculum is another way to support diversity. In a talk entitled "Leveraging Students' Cultural Competencies through Mathematical Modeling," Cynthia Anhalt described one way to do this by providing a glimpse of her work with in-service teachers. With collaborators, she designed the Mathematical Modeling in the Middle Grades (M^3) project, which brings culturally-relevant mathematical modeling curriculum to both teachers and students in southern rural Arizona near the U.S.-Mexico border. By offering professional development to these teachers, Anhalt and her colleagues support the teachers' use of engaging mathematical tasks that require students (many of whom are Hispanic/Latino) to model local community contexts and communicate their thoughts and results.

Edward Doolittle (First Nations University of Canada) described an additional approach to culturally-pertinent coursework in a talk entitled "Completing the Circle, Going Back to the Source: Indigenizing University Mathematics." Doolittle has been developing curriculum for his "Introductory Finite Mathematics" course in order to help his students find motivation, meaning, and relevancy in the mathematics that they do. For example, Doolittle makes use of the starblanket quilt pattern, much appreciated among many First Nations' students, to teach sequences and finite difference schemes. He uses the Mayan and Mohawk calendars to explain modular arithmetic and bases. These inclusions also allow students to reflect on cultural differences specific to First Nations people, such as the lesser importance of birthdays but the higher importance of rituals and their proceedings, length, and meanings.

Participating in a 30-minute training or modifying curriculum for cultural relevancy does not suddenly make you a diversity expert. In organizing this minisymposium, I hoped to give attendees ideas, enthusiasm, and confidence in the realization that there are things we can do to support diversity in our departments, thus making a difference for all of our students. Indeed, we can better support our underrepresented students, and set an example for all students about how to navigate situations involving people with different life experiences.

Many others are engaged in work related to diversity and inclusion, and more needs to be done in this area. Let this be a call for action that inspires us to include further sessions on diversity at the next SIAM Conference on Applied Mathematics Education and improve support of diversity in our departments.

Acknowledgments: The author would like to thank Cynthia Anhalt and Rachel Levy for their great suggestions on how to improve this article.

Rosalie Bélanger-Rioux is a preceptor and assistant director of undergraduate studies in the Mathematics Department at Harvard University, where she is involved in teaching, mentoring, and community building. She obtained her Ph.D. in applied mathematics from the Massachusetts Institute of Technology in 2014.

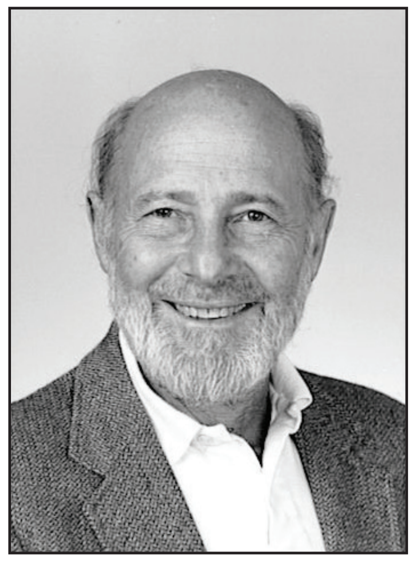
[1] Some presenter slides from the minisymposium are available at http://scholar.harvard.edu/rbr/presentations/diversity and http://scholar.harvard.edu/rbr/presentations/training

Obituaries

By Bernard J. Matkowsky

Joseph Bishop Keller, the foremost contemporary creator of mathematical techniques to solve problems in science and engineering, passed away on September 7, 2016. He earned this reputation by his outstanding research contributions to both mathematical methodology and a wide variety of applications. In addition, he taught and trained generations of applied mathematicians who form the "Keller School of Applied Mathematics." Through his own work, and that of his students and others with whom he interacted, he profoundly influenced the way that problems are formulated and solved mathematically. Keller combined unmatched creativity in developing mathematical methods with deep physical insight. He had an uncanny ability to describe real-world problems by simple yet realistic models, to solve the mathematical problem by sophisticated techniques, many of which he himself created, and then to explain the result and its consequences in simple terms. He was a master of asymptotics and in showing how to adapt ideas found useful in one area to others. His work is characterized by originality, depth, breadth, and elegance, and the results obtained have sustained importance. Due to space limitations, we shall only briefly describe certain highlights.

One of Keller's most outstanding contributions is the Geometrical Theory of Diffraction (GTD), which he originated for solving problems of wave propagation. He began thinking about such problems in work during World War II, on problems of sonar. GTD is an important extension of the Geometrical Theory of Optics (GTO), where wave propagation is described by rays. The extension overcomes difficulties which cannot account for phenomena such as diffraction, or the occurrence of signals where GTO predicts none. Keller developed a systematic way to treat high-frequency waves, and thus derived and solved the equations determining the rays, or paths along which signals propagate, as well as those governing how signals propagate along the rays. These include predictions of what happens as rays encounter obstacles or inhomogeneities of the medium in which they travel. Prior to Keller's work, only a few isolated problems were solved and understood, and there was no general theory for the solution of more complex and technologically important problems. Now there exist books devoted to Keller's theory, as engineers and scientists employ his theory to this day. Indeed, it is an indispensable tool for engineers and scientists working on radar, antenna design, and in general, on high-frequency systems in complicated environments. The impact of his work could be judged by attending a meeting of URSI, the international society devoted to radio science, where sessions were devoted to Keller's methods. This theory has been and is still applied to a variety of other problems in which signals are transmitted by waves, including acoustics, as in problems of sonar, and elastodynamics, as in quantitative nondestructive testing, and seismic exploration for oil, to name but a few. It is commonplace in all these fields to see articles which read, "we employ Keller's method to...".



Joseph Bishop Keller, 1923-2016. Photo credit: Stanford University.

Keller also showed that the methods developed for wave propagation were extendable to other classes of problems. Among these is his fundamental and penetrating work on semi-classical mechanics. In his work, Keller generalized work by Planck, Bohr, Sommerfeld, Wilson, Einstein and Brioullin to derive the correct quantization rules for non-separable systems, thus yielding results valid in any coordinate system. His results, referred to as the Einstein-Brioullin-Keller (EBK) quantization rules, are employed by many scientists. In his work on semiclassical quantization he introduced an important measure, the number of times a closed curve passes through a caustic surface, later generalized by Maslov and called the Keller-Maslov index. This too was subsequently extended by Keller to eigenvalue problems in bounded domains, not necessarily associated with quantum mechanics, but governed by general systems of partial differential equations.

Keller's work stimulated a vast literature in both the U.S. and abroad, not only in areas of science and engineering, where his methods and results are routinely employed, but in the mathematics community as well, where his work was taken up by pure mathematicians. For example, his work was the impetus for developments in the theory of Fourier Integral Operators and Lagrangian Manifolds.

In addition, he opened up directions of investigation by considering problem areas which were enthusiastically taken up by the research community. For example, his pioneering work on the evolution of singularities of nonlinear wave equations, as well as on bifurcation theory and nonlinear eigenvalue problems, to which scant attention was paid until the notes of his seminar appeared, is now one of the hottest topics of investigation by both pure and applied mathematicians alike.

Keller also considered problems of wave propagation through heterogeneous, turbulent, or random media involving the transmission of signals through media such as the atmosphere and oceans, in which fluctuations occur due to the irregular and fluctuating properties of the medium. He originated two methods, both widely used. The first is the Smoothing Method, for problems involving small amplitude variations, while the second is a Multiple Scale Method for rapidly varying coefficients. Thus, the second method deals with fluctuations which are not small in size, but rather small in scale. This theory, since taken up by others and now known as the Theory of Homogenization, has had volumes written on it. In each case, Keller showed how to systematically replace the fluctuating coefficients by effective coefficients, which are appropriate averages of the fluctuating coefficients. He then showed how to systematically derive effective equations for many problems, not necessarily associated with wave propagation. These include, e.g., problems of composite media and problems of determining the large-scale macroscopic behavior of a medium which exhibits small-scale microscopic heterogeneity. His work was characterized by a simple formulation which overcame the nonuniformities restricting earlier theories.

No stranger to national service, Keller worked on many problems related to nation-

See **Obituaries** *on page 11*

Q&A with NIH's Susan Gregurick

Susan Gregurick is the director of the National Institute of General Medical Sciences (NIGMS) Division of Biomedical Technology, Bioinformatics, and Computational Biology (BBCB) at the National Institutes of Health (NIH). She recently chatted with science writer and mathematician Analee Miranda about the NIH's mission and research focus, funding opportunities and programs available for applied mathematicians, career prospects at the NIH, and more.

What is the focus of your division?

The BBCB advances basic biomedical research. There are two branches within the division: the Bioinformatics and Computational Biology Branch and Biomedical Technology Branch. The long-term goals of the division are to leverage and advance data science and technologies to answer fundamental biological questions, to develop a computing infrastructure for the biomedical research community, and to promote and facilitate the development and use of new computational and experimental technologies in biomedical research by facilitating training opportunities.

What is the structure of each branch within your division?

We have two different branches, each headed by a branch chief. Under the branch chief are four health science administrators. The Biomedical Technology Branch focuses on the development and dissemination of new and novel technologies to advance research in biomedicine. The Bioinformatics and Computational Biology Branch focuses on the development of new and novel methods and computational tools that advance biomedicine.

What are the backgrounds of the program directors who work in your division? Do you have researchers employed in your division or in any of the supporting branches?

We are (completely, in our division and most of the NIGMS) extramural funding (resource) administrators, so we don't conduct our own research. We don't have researchers actively doing research in our division. Though all of our health science administrators were at one time researchers or research professors in academia, they gave up their own research programs to come and work as federal agents to administer investigator-initiated research grants.

NIGMS has a good record of collaborating with the National Science Foundation (NSF)/Division of Mathematical Sciences. How did this collaboration begin? What did NIGMS gain that might not have happened in the absence of the collaboration?

It's a longstanding collaboration that's over 12 years old and was built from the bottom up due to interest from the program staff both at the NIH and the NSF, who noticed the opportunity to bring mathematical and statistical science into biomedicine research. At that time, the NIH received very few applications in the mathematical sciences through its normal grant process, like the Parent R01, which focused on new and interesting mathematical research that would impact biomedicine. This program was started to address that gap. It's a real opportunity and also a wonderful collaboration between two different agencies. The NSF runs the application process – they issue the solicitation and receive all of the applications, then share these applications with the NIH. The applications are reviewed at the NSF under their review process. Once the review is concluded, both agencies get together and decide which of the meritorious and high-scoring applications should be funded by which agency. The NSF takes applications that are more focused in their mission area, and the NIH takes applications that focus much more on biomedicine. One of our goals is to create a pipeline for mathematical and statistical researchers to come to the NIH.

What kind of value do you think applied mathematicians bring that is vital to the success of NIH programs that you oversee?

They bring new concepts, new ideas, and new ways (insights) of looking at biomedical problems, a perspective that our researchers don't always have. That they offer fresh sets of eyes that are trained differently and have a different way of looking at problems is really what I think is pushing the innovation and impact of these programs. We're so very happy that they're contributing to our research. They push the boundaries.

What can you tell us about how an applied mathematician can request funding from your division? Is the process similar to that of the NSF? What does a mathematician need to know before submitting a proposal to the NIH?

I'd be more than happy to talk about the process of getting a grant in to the NIH. If you're not coming in through the NSF mathematical biology program, if you're just coming in through the Parent R01, the most important thing to do is identify the institute to which you want to apply. Our division tends to take a lot of mathematicians, so identifying a program officer and making a connection can help. You want to get a program officer to at least look at your rough idea. We *can't* look at your grant proposal but we're more than happy to give you feedback on your white paper. Identify your studies section very carefully. The studies sections at the NIH focus on different areas and have different compositions, different makeups, and different knowledge about mathematics and statistics, so you really have to make sure that your application is going to the right study section. Those are the two most important things that a potential principal investigator (PI) can do to ensure success here at the NIH. We are here to help and serve, but reaching us by phone is not always the best first choice; I would say send an email – we always respond to email! I'd also like to note that our joint mathematical biology program with the NSF has a sort of hybrid method (process) that's neither all NSF nor all NIH. That may be of particular interest.

What kind of positions in the NIH would be suitable for mathematicians? What are the requirements and how would they apply?

They can apply to work as postdoctoral fellows at the NIH. We have a Postdoctoral Research Associate (PRAT) Program for

See **Susan Gregurick** *on page 12*

SIAM Fellow Honored in China

SIAM Fellow John E. Hopcroft received China's Friendship Award in a ceremony at the Great Hall of the People in Beijing on September 29, 2016. Hopcroft is the IBM Professor of Engineering and Applied Mathematics in Computer Science at Cornell University. The Friendship Award is the highest honor given by the People's Republic of China to foreign experts who have made significant contributions to the country's economic and social progress.

"Talent is uniformly distributed around the world, and China has one fourth of the world's talent," Hopcroft said. "Helping China improve university education is an opportunity to impact the lives of 40 million students. To have this scale of impact is a once-in-a-lifetime opportunity, and the reason that I have spent the last five years helping China improve its university system."

SIAM Fellow John E. Hopcroft (left) receives the Friendship Award from Chinese Vice Premier Ma Kai (right) in September. Photo credit: Official Photo.

Pythagorean Theorem on Ice

By Mark Levi

In the spirit of winter fun, here is a short "skater's proof" of the Pythagorean theorem. Starting in the corner O of a skating rink with the walls along the x and y axes, wearing perfectly slippery shoes on the perfectly slippery ice, I push away from the y-wall, acquiring speed a in the x-direction (see Figure 1). Next, I push away from the x-wall, gaining speed b in the y-direction. With the first push I acquired kinetic energy $ma^2/2$, and with the second push I added $mb^2/2$ to my kinetic energy. Indeed, the fact that I was sliding along the x-wall is irrelevant, because my gloves are perfectly slippery; it feels the same as if the wall were not sliding by at all, just like during the first push.[1] After the two pushes, my speed c is the hypotenuse of the velocity triangle, and thus my kinetic energy is $mc^2/2$. But this energy is the accumulation of the two previous contributions:

MATHEMATICAL CURIOSITIES
By Mark Levi

$$\frac{mc^2}{2} = \frac{ma^2}{2} + \frac{mb^2}{2},$$

implying $a^2 + b^2 = c^2$. All this can be summarized by saying that the energies add as scalars, while the velocities add as vectors. Of course, the above is not meant as a rigorous proof and is rather an interpretation, or a physical incarnation, of the Pythagorean theorem.

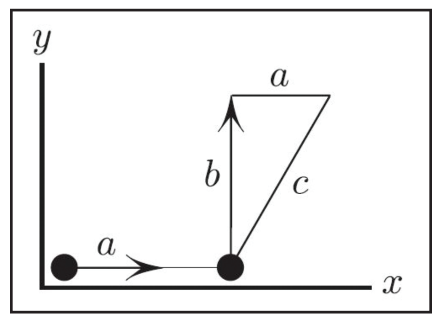


Figure 1. *A kinetic energy "proof" of the Pythagorean theorem.*

The figure in this article was provided by the author.

Mark Levi (levi@math.psu.edu) is a professor of mathematics at the Pennsylvania State University.

[1] it is here that the orthogonality of the walls is used.

Professional Opportunities and Announcements

Send copy for classified advertisements and announcements to: ***marketing@siam.org****; For details visit* ***www.siam.org/advertising****.*

*Students (and others) in search of information about careers in the mathematical sciences can click on "Careers and Jobs" at the SIAM website (***www.siam.org****) or proceed directly to* ***www.siam.org/careers****.*

NRC Postdoctoral Research Positions

NIST — INFORMATION TECHNOLOGY LABORATORY

The Applied and Computational Mathematics Division (ACMD) of the National Institute of Standards and Technology (NIST) invites applications for two-year NRC postdoctoral research positions at NIST Laboratories in Gaithersburg, Maryland, and Boulder, Colorado. NIST is a Federal government research laboratory specializing in measurement science. ACMD consists of some 47 full-time professional staff, along with part-time faculty appointees and guest researchers. Staff members engage in collaborative research with scientists throughout NIST, providing expertise in applied mathematics, mathematical modeling, and computational science and engineering.

Research areas of interest include combinatorial and discrete algorithms, computational materials science, computational fluid dynamics, computational electromagnetics, computational biology, orthogonal polynomials and special functions, applied optimization and simulation, combinatorial software testing, data mining, immersive visualization, parallel and distributed algorithms, quantum information science, and statistics for quantum systems.

Of particular interest are candidates whose interests also overlap with the thrusts of recent Federal initiatives, including the National Strategic Computing Initiative (https://www.whitehouse.gov/blog/2015/07/29/advancing-us-leadership-high-performance-computing) and the Materials Genome Initiative (https://mgi.nist.gov), especially in the area of uncertainty quantification for modeling and simulation.

Candidates and their research proposals are evaluated in a competitive process managed by the National Research Council (NRC) Associateship Programs. The current stipend is $67,588 per year. For further details, see http://www.nist.gov/itl/math/mcsd-postdoctoral-opportunities.cfm. Application deadlines are February 1, 2017 and August 1, 2017. Appointments commence within one year of selection. For questions, contact Tim Burns, burns@nist.gov.

NIST is an equal opportunity/affirmative action employer. The NRC Associateship Program at NIST is restricted to US citizens.

National Science Foundation Call for Proposals

Algorithms in Threat Detection (ATD)

The Algorithms for Threat Detection (ATD) program will support research projects to develop the next generation of mathematical and statistical algorithms for analysis of large spatiotemporal datasets with application to quantitative models of human dynamics. The program is a partnership between the Division of Mathematical Sciences (DMS) at the National Science Foundation (NSF) and the National Geospatial Intelligence Agency (NGA).

Deadline for proposals is **February 21, 2017**.

More information can be found at https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503427.

National Science Foundation Call for Proposals

Algorithms for Modern Power Systems (AMPS)

The Algorithms for Modern Power Systems (AMPS) program will support research projects to develop the next generation of mathematical and statistical algorithms for improvement of the security, reliability, and efficiency of the modern power grid. The program is a partnership between the Division of Mathematical Sciences (DMS) at the National Science Foundation (NSF) and the Office of Electricity Delivery & Energy Reliability (OE) at the U.S. Department of Energy (DOE).

Deadline for proposals is **February 13, 2017**.

More information can be found at https://www.nsf.gov/pubs/2017/nsf17521/nsf17521.htm.

samsi
NSF•Duke•NCSU•UNC

SAMSI SEEKING NEW DEPUTY DIRECTOR

The Statistical and Applied Mathematical Sciences Institute (SAMSI) invites applications for the position of Deputy Director for a term of two years **beginning July 1, 2017**.

The Deputy Director will be a distinguished researcher who will provide scientific direction to the institute and oversight of the SAMSI grant, and who will work closely with the Director on all aspects of the institute's oversight and program activities. The Deputy Director will also be strongly encouraged to pursue personal research in conjunction with the SAMSI programs or independently.

Together with the Director, the Deputy Director forms the executive side of the SAMSI Directorate whom are responsible for the administration of programs, human resources, financial operation and infrastructure. Together with the other members of the Directorate, they also share the responsibilities of the selection, development and implementation of SAMSI programs.

The appointment will be made as a member of the research faculty at North Carolina State University.

Candidate must have a minimum of a Ph.D. in Mathematics or Statistics or equivalent.

Qualified candidates should be mathematicians or statisticians with excellent management skills and research record. Proven administrative and operational experience is an asset. In addition, the successful candidate will demonstrate a strong interest in further developing and expanding the mission of the institute.

Additional information and a link to **N.C. State University's Job site** for submitting applications may be found at: **https://jobs.ncsu.edu/postings/76044**.

Candidates are asked to attach a current curriculum vitae, letter of application, and contact information for three professional references. Informal inquiries may be addressed to **Richard Smith**, **Director of SAMSI**, rls@samsi.info. **Review of applications will begin February 2, 2017 and will continue until position is filled.**

Individuals with disabilities requiring disability-related accommodations in the application and interview process, please call **919-515-3148**. Final candidates are subject to criminal & sex offender background checks. Some vacancies also require credit or motor vehicle checks. If highest degree is from an institution outside of the U.S., final candidates are required to have their degree verified at www.wes.org. Degree must be obtained prior to start date.

NC State University participates in E-Verify. Federal law requires all employers to verify the identity and employment eligibility of all persons hired to work in the United States.

Obituaries

Continued from page 8

al security and served on various advisory boards, national panels, and committees. After his work on sonar for the Columbia University Division of War Research, he worked on problems of underwater explosions in order to predict the shock wave and water waves to be expected at the Bikini atomic bomb tests. At the time, there was concern of producing a tsunami which might devastate Japan and other Pacific countries. His analysis showed there was no danger. He also spent time at Argonne and Los Alamos National Laboratories, studying hydrogen bomb explosions. In the early 1950s he served, with Von Neumann, on a Committee on Underwater Atomic Bombs for the Air Force to consider the effects of A-bomb explosions on ships and submarines, and headed another project on A-bomb explosions. During the late 1960s he was a member of JASON, a group of high-level consultants to the Defense Department and other governmental agencies on scientific and technical matters. Furthermore, he served as consultant to the Air Force Special Weapons Project, U.S. Naval Air Development Center, U.S. Army Chemical Corps, and Argonne, Brookhaven, and Los Alamos National Laboratories.

In addition to his important and prolific research, Keller was a teacher and expositor par excellence. He twice received the MAA's Lester R. Ford Award for outstanding expository writing. He received awards from all three major U.S. mathematical societies, from various engineering societies, as well as from national honorary societies, both in the U.S. and abroad. The 60 Ph.D. students and numerous postdoctoral associates whom he has trained, now successful applied mathematicians in their own right, further attest to the impact that Keller had. In short, Keller was one of the most prolific and important investigators and educators of our time.

Finally, there is Joe Keller the man. Countless numbers of mathematicians, engineers and scientists have come to him through the years to benefit from his acumen and understanding. To each he listened patiently, contributed helpful insights, and offered words of advice and encouragement. For each he was simply "Joe," teacher, colleague, and friend. The world has lost a giant. He will be sorely missed, though his legacy endures.

Bernard J. Matkowsky is the John Evans Professor of Engineering Sciences and Applied Mathematics, Mathematics, and (by courtesy) Mechanical Engineering at Northwestern University.

Machine Learning's Impact on Global Public Health

Computational Approaches to Verbal Autopsies

By Paul Davis

How bad is malaria in sub-Saharan Africa? Whom does malaria disable or kill? Where does it strike? Are its dangers rising or waning? Which threat is worse, malaria or, say, road injuries?

The Global Burden of Disease (GBD) study offers answers to these kinds of questions to help public health officials choose between distributing mosquito nets and installing traffic signals. Through a minitutorial at the 2016 SIAM Annual Meeting, held in Boston, MA, last July, Abraham Flaxman (Institute for Health Metrics and Evaluation at the University of Washington) introduced applied mathematicians to this important health policy tool. Then he led his audience on a deeper, hands-on dive into one of the study's computationally sophisticated components: verbal autopsies.

For your own version of Flaxman's overview of the GBD study, go to the GBD Compare page on the Institute for Health Metrics and Evaluation site,[1] or see Figure 1. These so-called tree maps display by country (though the entire world appears in Figure 1) the years of life lost to disability and death from infectious disease (red), noncommunicable disease (blue), and injury (green). Technically, these maps display disability-adjusted life years (DALYs), a construct that captures both death and disability in proportion.

Users of these online tree maps and other visuals can track changes over time in one country or region or disease; e.g., the peak in deaths and disability due to AIDS in Africa is painfully clear. Or they can track risk factors—matters that societal health policy can affect—like obesity, which is indicated by a high Body Mass Index.

The data underlying these visual displays is an account of who died from which cause. It is seemingly simple to gather, if you knew who died in the region of interest, say sub-Saharan Africa. And if you could see every death certificate. Assuming none of the death certificate entries were mistaken, perhaps confusing a risk factor (hypertension) with a cause (aneurysm). And presuming each certificate listed but a single cause of death, a supposition any clinician would reject out of hand.

Or perhaps you could succeed by changing tack to extract causes of death from hospital data, assuming it is valid, and somehow extend those causes to the population that died outside of hospitals throughout sub-Saharan Africa.

In the spirit of a tutorial, Flaxman paused at this discouraging juncture to "pair and share," promoting engagement by letting audience members chat with those around them and come up with their own responses to this data challenge. Once everyone was invested in the problem, he explained verbal autopsies, the process now in use to determine cause of death.

In a verbal autopsy, a trained interviewer asks those who were with the deceased a series of structured questions, such as "Did your father have a fever during his last illness?" From the resulting narrative of the final illness, a health professional with experience in the specified location and culture can identify a cause of death.

Flaxman focuses on the use of machine learning to assign a cause of death in place of that experienced health professional. The balance of his tutorial introduced some tools of the trade via an online computational notebook, before moving from those immediate computational experiences to more subtle questions.

Which quantity should we optimize to accurately and usefully interpret the narrative of a verbal autopsy? How do you test the machine learning tool – by holding back an expensive subset of this "big cost data?" Are machine learning tools almost too flexible for useful generalizations in this setting?

Flaxman confesses to being "obsessed with reproducibility" in his decisions about machine learning training and testing. What metric of accuracy should we optimize to avoid being "wrong about half the time" when the tool is used out of sample? His choice on that front is the "cause-specific mortality fraction," a kind of relative error not for the individual causes of death but for the overall error in the fraction of the deceased assigned each particular cause of death.

Flaxman seeks to "explain why," in the sense of identifying particular answers to the verbal autopsy's structured questions that led the interviewer to select certain causes of death. He suggests that "understanding errors builds trust" in the computational methods, certainly a valuable perspective in such a diverse disciplinary environment as the GBD study. An immediate practical challenge is achieving similar accuracy when identifying causes of death using fewer questions and therefore shorter, less expensive interviews, or so-called "data-driven item reduction."

Flaxman's insightful tutorial exposed a hard, practical, and immensely important contribution of machine learning to global public health. He cleverly let his audience dip their toes into puddles of questions surrounding formulation and implementation. And his litany of tough queries revealed that many significant challenges remain.

Links[2] to a video and PDF of Flaxman's presentation at the 2016 SIAM Annual Meeting offer further information. Students can also learn about post-baccalaureate fellowships[3] offered by the Institute for Health Metrics and Evaluation.

Paul Davis is professor emeritus of mathematical sciences at Worcester Polytechnic Institute.

[1] http://www.healthdata.org/data-visualization/gbd-compare

[2] https://www.pathlms.com/siam/courses/3028/sections/4133

[3] http://www.healthdata.org/post-bachelor-fellowship

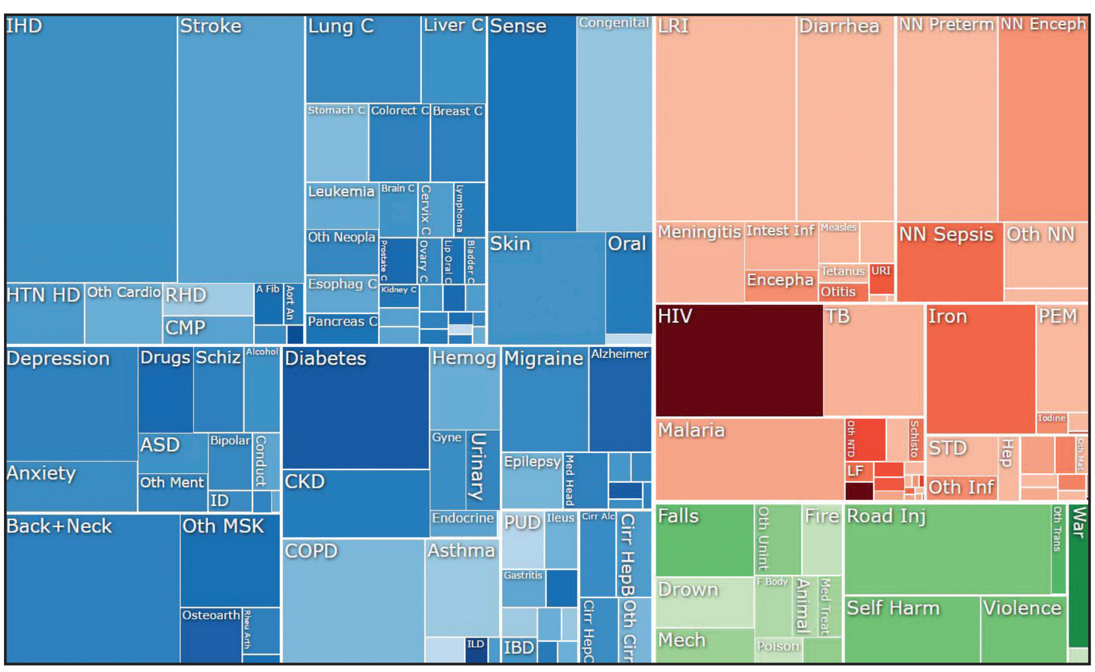


Figure 1. *This tree map displays the relative contribution worldwide to mortality and disability in 2015 from infectious disease (red), noncommunicable disease (blue), and injury (green). Image credit: Institute for Health Metrics and Evaluation (IHME). GBD Compare. Seattle, WA: IHME, University of Washington, 2016. Available from http://vizhub.healthdata.org/gbd-compare. (Accessed November 15, 2016).*

Susan Gregurick

Continued from page 9

postdoctoral fellows to partner recent Ph.D. graduates with intramural researchers at the NIH. We also get a number of fellows from the American Association for the Advancement of Science (AAAS) who do policy and analysis work. It's a postdoctoral-type fellowship where they are not necessarily doing research, but doing much more programmatic work. That's a great career path for some. In fact, we have an AAAS Fellow now who is a cosmologist and string theorist. He has no background at all in biomedical science or medicine.

How has your division handled the funding constraints brought about by Congress in recent years? Now that Congress has increased funding and may increase it again over the next few years, what do you see as the priorities for the NIH moving forward?

It's tough whenever we get funding constraints by Congress. Our mission is really to prioritize and fund the most impactful and meritorious science that we receive through our grant application process. You'll see that our success rate (for proposals) has dipped in prior years, and that was troubling to many people. However, with some really innovative new ideas as well as increased funding from Congress, we now have a success rate for proposals at a near-record high. In Fiscal Year 2015 (FY-15) I think it was at 28%, and we've probably been able to maintain that in FY-16.

I also want to tell you about one innovative program that my director Dr. Jon Lorsch started just last year as a pilot project, which we will soon implement as a full-scale program. It is called the Maximizing Investigators' Research Award (MIRA) program, and is applicable to anybody in biomedical research that falls within the mission of NIGMS – mathematicians and statisticians doing work that is relevant to biomedicine. It is a sustaining grant, starting at usually $250K/year and given to any investigator to fund all of his or her work in an NIGMS-related laboratory. It offers a longer preparative time and requires a slightly shorter, more visionary proposal. We hope to sustain participants' labs for many years, so they don't have to continually write more and more R01s to have a career in science, given that they make good progress on their research. This is just one more avenue for folks who are part of the SIAM community to come to the NIGMS and contribute to our mission. We especially encourage early-stage investigators – someone within ten years of their Ph.D. or an assistant professor at a university or college.

Since our mathematical biology program is teamed with the NSF, we make sure that we are pulling mathematicians and statisticians to our research areas. The program did not sustain any cuts during this funding period but there were cuts in other areas, so in some sense they (the biomathematicians) were buffered a little because this is a high-priority program for us.

Do you have international collaborations within your group? Are there satellite offices of the NIH in other countries? Are there funding opportunities for international students (or researchers) who are interested in working with the NIH?

If there's a collaboration that a PI is bringing through their grant, then we are certainly happy to fund those (informal international) collaborations. But Joint Programs, some that are international, are helped through the Fogarty International Center. To my knowledge, we don't have any international programs, but we certainly watch international sciences through collaborative science proposals.

Are there internships available for applied math students who are interested in biological applications of mathematics?

We don't have junior program officer positions, but we do, through the AAAS program, hire a number of people from the Ph.D. to postdoctoral levels; after the postdoctoral period, we do hire them in our Office of Program Planning, Analysis and Evaluations. However, almost all of our program officers have been academic researchers and have either received tenure or worked in national laboratory settings and been promoted to the equivalent of a tenured position. Although it's not always true, we want to hire people who understand what it is like to be a PI and to conduct a research program, manage a budget, and manage graduate students.

Susan Gregurick, Ph.D., is director of the National Institute of General Medical Sciences (NIGMS) Division of Biomedical Technology, Bioinformatics, and Computational Biology (BBCB). In this capacity, she oversees programs that join biology with the computer sciences, engineering, mathematics, and physics. Prior to joining NIGMS, she served as acting director of the Biological Systems Science Division at the Department of Energy (DOE). She also developed and managed the DOE's Systems Biology Knowledgebase. You can contact her at greguricksk@mail.nih.gov.

Analee Miranda is a freelance curriculum writer and mathematician. She is a subject matter expert on the radar scattering effects of UHF/VHF/HF radar on humans. Her current research interests include discovering new medical applications of radar technology.