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

siam news

Volume 49/ Number 2
March 2016

See Light Move: Compressed Sensing and the World's Fastest 2-D Camera

By Hans Kaper

It may take one, ten, or even a hundred years, but beautiful ideas always find their way to applications.

In an ordinary digital camera, a large amount of information is collected by the camera's sensor. This information is immediately processed by compression software, and a much smaller amount is then stored. This process takes time, contributing largely to what is called "shutter lag." Is it possible to collect only what is needed and thus reduce or even eliminate shutter lag? Thanks to some clever physics and compressed sensing, the answer is "yes."

The mathematical concepts underlying *compressed* or *compressive sensing* date back to the 1970s, when the merits of the ℓ^1 norm were recognized in convex programming. The theory developed rapidly in the 1990s, and a breakthrough occurred in the early years of the 21st century with the work of Candès, Tao, Romberg, and Donoho [2, 3, 5]. Collectively, they realized that one can design an efficient sensing or sampling protocol by capturing useful information in a sparse signal and compressing it into a small amount of data.

These beautiful mathematical ideas have found an application in the world's fastest 2-D camera. By adding a digital micromirror device for spatial encoding and applying reconstruction algorithms from compressed sensing, Lihong Wang and colleagues at Washington University in St. Louis transformed a traditional one-dimensional streak camera into a 2-D ultrafast imaging camera. Now, for the first time, humans can see the movement of light on the fly (see Figure 1).

Compressed Sensing

Compressed sensing is based on the observation that most signals or images contain only a small amount of crucial information. By strategically discarding unnecessary data, one can reduce the image's size while retaining sufficient information to faithfully reconstruct the original. The full image is then reconstructed from the encoded data using linear programming algorithms. To make this possible, compressed sensing relies on two principles: *sparsity* and *incoherence*. Sparsity implies that the image has a concise representation when expressed in the proper basis, while incoherence indicates that unlike the image of interest, the sampling/sensing waveforms have an extremely dense representation in the sparsifying basis. Thus, compressed sensing is a simple and efficient signal acquisition protocol that samples—in a signal-independent fashion—at a low rate and later uses computational power for reconstruction from what appears to be an incomplete set of measurements.

Instead of collecting all of the information about a signal or image in pixels in a large n-dimensional vector x, compressed sensing collects information in a much smaller k-dimensional vector $y = Ax$. Here, A is a (rectangular) matrix that represents the measurement process. If the signal is sparse, k need not be very large to recover the original signal x from y; in practice, it usually needs to be no more than four times the sparsity level of x.

The underdetermined system $Ax = y$ can be solved in a remarkably efficient way by minimizing the ℓ^1 norm $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ of x, subject to the constraint $y = Ax$. This convex problem can be converted to a linear

See **Compressed Sensing** *on page 3*

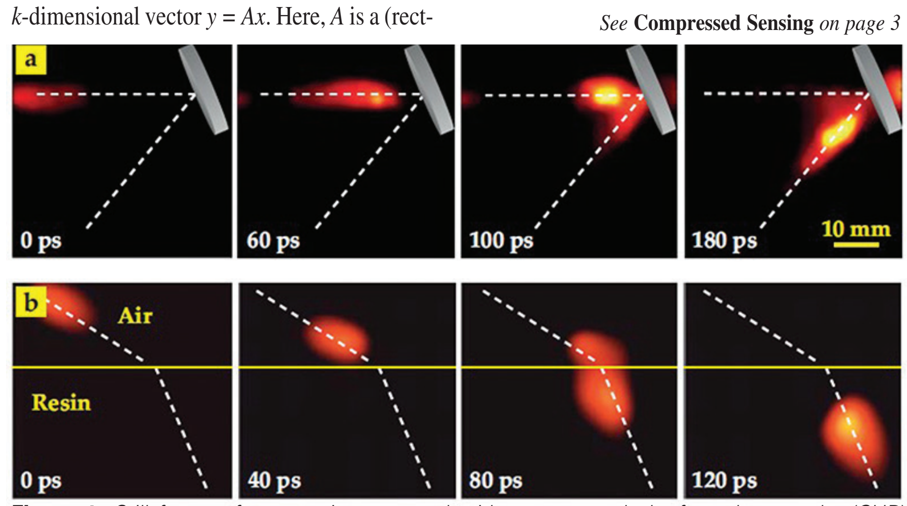


Figure 1. *Still frames from movies captured with compressed ultrafast photography (CUP) shows a light pulse (a) bouncing off a mirror and (b) refracting at an air-resin interface.*

Computational Modeling of Convection in the Earth's Mantle

By Wolfgang Bangerth, Juliane Dannberg, Rene Gassmöller, and Timo Heister

Louis Moresi's article[1] in the December issue of *SIAM News* prompts readers to take a deeper look at models of Earth – deeper into Earth, that is, and more specifically into Earth's mantle: the region between the solid crust and the metallic core.

The mantle makes up about 84% of Earth by volume and extends to a depth of 2,890 km. At its bottom, it is believed to reach temperatures of about 3,300–4,400 K. Yet because of pressure from overlying material, the mantle consists of solid rock (with few localized exceptions of partially molten material), as evidenced by its ability to transmit seismic shear waves. However, on sufficiently large time scales it behaves like a vigorously convecting fluid, cooled by the cold surface above and heated by both the hot core below and radioactive decay within (see Figure 1). Thus, one could view plate tectonics simply as the surface expression of convection cells. Within the mantle, average velocities are on the order of cm/year.

At first glance there seem to be only a few visible expressions of mantle convection. However, most volcanism on Earth's surface is closely related to the distribution and advection of thermal and chemical heterogeneities in the mantle and at mid-ocean ridges, subduction zones, and oceanic islands, for example. Furthermore, the importance of the mantle lies in its interactions with other parts of Earth:

• When hot material rising in the mantle approaches the surface, it causes massive melting that leads to large volcanic eruptions. The volume of erupted material can exceed 10^6 km^3 in just a few million years, leading to the formation of large volcanic provinces with areal extensions greater than 10^6 km^2. Due to the released gases, these events influence the global climate and may have led to mass extinction events.

• Cooling by the mantle drives convection in Earth's core, which generates Earth's magnetic field. A more viscous mantle would effectively insulate the core, and a less viscous mantle would have led to the core's crystallization long ago; both scenarios would leave us without the magnetic field that offers protection from the harsh radiation of space and keeps solar wind from driving away the atmosphere.

• By ingesting water and sediment-laden subducting oceanic plates—and later releasing material again through volcanic outgassing—the mantle acts as a giant reservoir for water and carbon, thus creating a carbon cycle on time scales of tens of millions of years.

For these reasons, understanding the dynamics of convection and the conditions under which it happens have long been important in the geosciences. More recently, a fair number of computational and applied mathematicians have also entered the field, as mantle convection presents an attractive, accessible set of problems that require large-scale computations, are tough on linear and nonlinear solvers, yet still fall within a solvable range.

Many aspects of mantle convection can already be understood qualitatively by considering a set of equations in which fluid flow is slow and driven by density differences that result from temperature variation. The Stokes equations allow for adequate modeling of velocity and pressure:

$$-\nabla\cdot(2\eta\varepsilon(u)) + \nabla p = \rho g,$$

$$\nabla\cdot u = 0,$$

with viscosity η (ranging from 10^{18} to 10^{23} Pa·s; for comparison, the viscosity of water is about 10^{-3} Pa·s; that of air 10^{-5} Pa·s; and that of honey ~10 Pa·s), density ρ, and gravity g. An advection diffusion equation models temperature:

$$\frac{\partial T}{\partial t} + u\cdot\nabla T - \nabla\cdot\kappa\nabla T = Q.$$

This must be augmented with appropriate boundary conditions for the velocity and temperature, such as tangential or zero velocity around the boundary, and "hot" at the bottom and "cold" at the top of the domain. In the simplest description, the density depends linearly on the temperature, $\rho(T) = \rho_{ref} - \alpha(T - T_{ref})$, where α is the thermal expansion coefficient. With realistically small values for the thermal diffusivity κ, such a model already (correctly) predicts that heat is transported from the core to the surface primarily through "blobs," or sheets

See **Earth's Mantle** *on page 3*

[1] Moresi, Louis. (2015, December). Computational Plate Tectonics and the Geological Record in the Continents. *SIAM News*, *48*(10), 1 & 3.

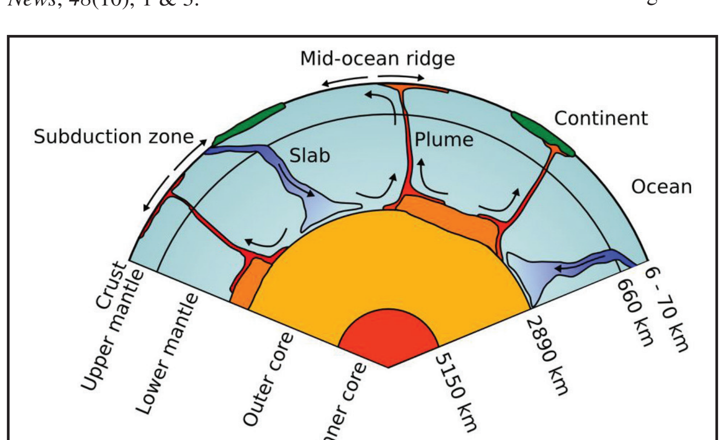
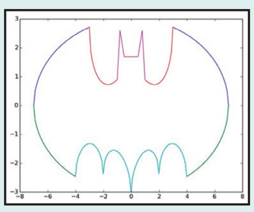


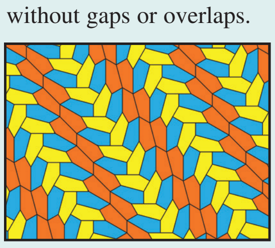
Figure 1. *Schematic of convection in Earth's mantle, including cold downwellings (blue), hot upwellings (red), and strong chemical heterogeneities (orange, green), all of which influence the material properties of the rocks.*

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 49/ Number 2/ March 2016





Large-Scale Inversion in Exploration Seismology

By Tristan van Leeuwen

Seismic data offers a rich source of information about the subsurface of the earth. By studying its dynamic and kinematic properties, researchers can infer large-scale variations as well as rock properties on a local scale. Seismic measurements for exploration purposes are typically acquired by placing receivers (geophones) on the surface and detonating an explosive source, as seen in Figure 1. This procedure is repeated for various locations, resulting in a large volume of data. This is a typical *multi-experiment* setting, meaning multiple data sets are collected for a single set of parameters. The rock properties are parametrized in the subsurface by m and the experiment is simulated by solving a linear wave equation $\mathcal{L}[m]u_i = q_i$, where $i = 0,1,\ldots,\ k$ is the experiment index, q_i represents the explosive source, and $\mathcal{L}$ is a differential operator.

The introduction of a linear operator $\mathcal{P}$ that maps the solution u_i to the measurements formally poses the inverse problem as follows: *For given measurements* d_i, *determine the coefficients m and solutions* u_i *such that* $\mathcal{P}u_i \approx d_i$ *and* $\mathcal{L}(m)u_i = q_i$ *for* $i = 1,2,\ldots k$.

Numerically solving the PDEs readily eliminates the u_i and obtains a high-dimensional (m may represent up to 10^9 parameters) nonlinear least-squares problem with k terms:

$$\min_m \sum_{i=1}^{k} \|\mathcal{P}u_i - d_i\|_2^2,$$

where $\mathcal{L}(m)u_i = q_i$ [5]. In principle, any black-box optimization method can be used to solve the resulting optimization problem. Due to the computational cost and severe nonlinearity, however, the seismic problem is not amenable to a black-box approach. The key to developing a better approach is considering the interplay between the formulation, the optimization algorithm, the multi-experiment nature of the data, and the means of (numerically) solving the wave equation. These aspects are traditionally different disciplines' areas of expertise (e.g., statistics, computer science, machine learning, and numerical analysis), making this a very exciting problem for multidisciplinary research.

The leading computational cost lies in solving the wave equation for all k experiments, where k is potentially very large (easily $k \sim 10^6$). Thus, one can only perform a few iterations to obtain an approximate solution of the optimization problem. Additionally, the severely nonlinear relation between the parameters and the data requires a very good initial parameter estimate. If the initial guess is not 'close' to the true parameters in some sense, the optimization may converge to a local minimum. Failure to find a global minimum is often very hard to detect. The industry therefore spends a considerable amount of time and effort constructing a suitable initial estimate and performing subsequent quality control, both of which involve much specialized manual interference.

An ideal situation would involve running an inversion multiple times from a suite of initial guesses and quantifying the uncertainty of the final result. While mathematical techniques to perform such uncertainty quantification for inverse problems are well established, they often rely on some form of Monte Carlo (MC) sampling. However, the high dimensionality of the problem at hand and the computational cost involved in running even a single simulation prohibit the use of such techniques. My recent research aims at tackling this challenge on multiple fronts:

- Reformulation of the conventional least-squares problem to one that is less nonlinear in the parameters, making the approach less sensitive to the quality of the initial guess [4, 7].
- Reduction of the dimensionality of the data, replacing the full data set with k terms by a subsampled dataset with $k' \ll k$ terms [6].
- Better quantification of the uncertainty by estimating statistics of the noise and other auxiliary parameters [1].

Reformulations

Many reformulations of the seismic inverse problem have been proposed over the years. One class of reformulations uses a different distance metric to measure the difference between the observed and simulated data. This is very useful when certain features of the data are of primary interest, or when the data contains large outliers.

See **Exploration Seismology** *on page 4*

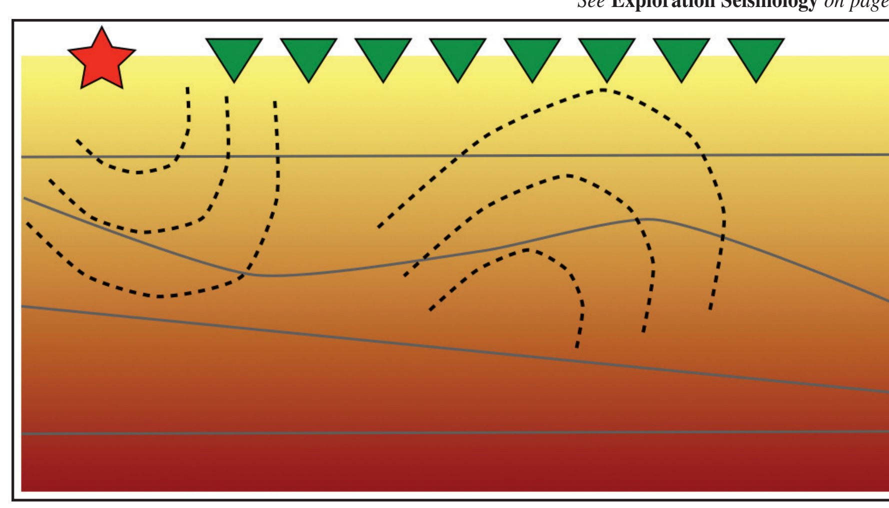


Figure 1. *Schematic depiction of the acquisition process. The seismic source is indicated with a ★ while the receivers are indicated with a ▼.*

New Initiative Focuses on Computational Science and Advanced Computing

With the launch of the federal National Strategic Computing Initiative (NSCI), the Obama Administration and federal agencies have thrust computational science and advanced computing into the spotlight. The initiative, which should be of interest to the SIAM community, aims to create a cohesive, strategic vision in High Performance Computing (HPC) for the United States. As agencies seek to implement this vision, applied math and computational science will play a key role in ensuring that the U.S. realizes the full benefit of HPC systems. Three primary agencies are engaged in the initiative: the Department of Energy (DOE), the National Science Foundation (NSF), and the Department of Defense (DOD).

The NSCI aims to combine the strengths of computers focused on simulation/modeling and computers focused on managing large amounts of data. The goal is to create new platforms, keep the U.S. at the forefront of HPC capabilities by achieving exascale computing, create systems that are easier to program to encourage widespread use, make HPC readily available through deployment and education, and conduct fundamental research to establish hardware technology for future HPC systems. For example, the NSF is planning investments in software, technology, and people. These investments primarily focus on increasing synergies between modeling/simulation and data analytic computing, increasing the capacity and capability of a national HPC ecosystem, and researching a viable path forward for future HPC systems post Moore's law. The DOE will continue its pursuit of an exascale system with accompanying applied math investments. The DOD will focus on data analytic computing related to its mission. As these agencies develop their implementation plans, new opportunities are likely to emerge for the SIAM community.

Given the growing nature of this initiative and the importance of modeling and other mathematical challenges to its success, the SIAM community is encouraged to participate in workshops, symposia, conferences, and Proposers Days to shape and compete for new opportunities arising from the NSF, the DOE, and the DOD. For instance, the White House held a workshop in October to discuss key challenges and opportunities related to the initiative, foster collaboration, and communicate long-term plans for the NSCI. Multiple participants at the workshop emphasized the need for research on algorithms, cited applied mathematics challenges, or noted the importance of computational science. The White House is currently planning follow-on workshops on particular challenge areas. Now is a great time to engage, bringing the expertise of the mathematics community to ensure that the U.S. will continue its leadership in HPC. *– Miriam Quintal, Lewis–Burke Associates, LLC.*

siam news

ISSN 1557–9573. Copyright 2016, all rights reserved, by the Society for Industrial and Applied Mathematics, SIAM, 3600 Market Street, 6th Floor, Philadelphia, PA 19104–2688; (215) 382–9800; siam@siam.org. To be published ten times in 2015: January/February, March, April, May, June, July/August, September, October, November, and December. The material published herein is not endorsed by SIAM, nor is it intended to reflect SIAM's opinion. The editors reserve the right to select and edit all material submitted for publication.

Advertisers: For display advertising rates and information, contact Kristin O'Neill at marketing@siam.org.

One-year subscription (nonmembers): Electronic-only subscription is free. $73.00 subscription rate worldwide for print copies. SIAM members and subscribers should allow 8 weeks for an address change to be effected. Change of address notice should include old and new addresses with zip codes. Please request address change only if it will last 6 months or more.

Printed in the USA.

siam *is a registered trademark.*

Earth's Mantle

Continued from page 1

of hot material that rise up from the core-mantle boundary and sheets of cold material that well down from the surface (see Figure 2). We can identify these hot blobs with mantle plumes believed to give rise to Earth's hot spots, and the cold sheets with subducting oceanic plates.

In this simple model, dynamics are easy to understand when considering the Rayleigh number $Ra = \alpha\, \Delta T g D^3 / (\eta\, \kappa)$, where D is a length scale of the domain and ΔT is the temperature difference between the surface and the core-mantle boundary; the larger the Rayleigh number, the smaller the features of the flow field. This already gives rise to a formidable computational challenge: with realistic values for the material parameters in the above equations, Earth's flow features can be as small as a few kilometers across. With a volume of around 10^{12} km^3, a finite element discretization using uniform meshes requires about a billion cells and a few 10^{10} unknowns to achieve appreciable accuracy. The current generation of codes can reach into this range, either through highly-tuned numerics or the use of adaptive mesh refinement. Some of the largest implicit finite element computations have indeed been performed on this problem [1, 2, 3, 5, 6, 7]. The most recent Gordon Bell prize was also awarded for the solution of this system on up to 500,000 cores [4].

A separate, equally-difficult challenge arises from the fact that realistic materials do not behave as outlined above. Rather than expand linearly with temperature, rocks undergo phase changes where density varies both continuously and discontinuously as a function of temperature and pressure. The same is true for viscosity, which may vary by many orders of magnitude even over small distances where hot and cold materials come together, such as when cold oceanic slabs subduct into the hot mantle. Viscosity also depends strongly and nonlinearly on stress, grain size, water content, and a number of other quantities. Finally, when considering the entire mantle, a mass conservation equation, $\nabla \cdot (\rho u) = 0$, must replace the above incompressibility equation.

Cumulatively, compressibility and strong nonlinearities complicate the design of efficient and accurate solver strategies. The required nonlinear iterations also make the solution very expensive computationally. While the nonlinearity can be treated efficiently in a time-stepping scheme, solving the first time step self-consistently can become a significant challenge.

Equally difficult are the large jumps in viscosity that result from strong temperature gradients that often cannot be resolved adequately by the mesh. Such "essentially discontinuous" viscosity fields cause large discretization errors and pose enormous challenges to the design of linear solvers and preconditioners. Recent experiments show that appropriately averaging material parameters on each cell, without reducing the overall convergence order, can significantly reduce these discretization errors. This also vastly improves the efficiency of linear solvers, sometimes reducing the time to solution by a factor of ten on complex models.

Much more progress in all areas of computational mathematics—on discretizations, nonlinear and linear solvers, preconditioners, and parallel algorithms—is necessary to make the solution of complex models in mantle convection fast and routine. However, many of the most widely-used codes are open source and well-documented, such as Citcom,[2] or our own contribution, ASPECT.[3] Specifically, ASPECT is built as a modular platform that allows mathematical and computational scientists to test new discretizations and solvers on realistic problems, and geoscientists to develop and test new model descriptions. The Computational Infrastructure for Geodynamics[4] has been collecting and curating these and other codes to facilitate both geodynamical research as well as experimentation with new numerical methods.

[2] https://geodynamics.org/cig/software/citcoms/
[3] http://aspect.dealii.org
[4] http://www.geodynamics.org

References:

[1] Burstedde, C., Stadler, G., Alisic, L., Wilcox, L. C., Tan, E., Gurnis, M., & Ghattas, O. (2013). Large-scale adaptive mantle convection simulation. *Geophysical Journal International*, *192*(3), 889-906.

[2] Gmeiner, B., Huber, M., John, L., Rüde, U., & Wohlmuth, B. (2015). A quantitative performance analysis for Stokes solvers at the extreme scale. *Cornell University Library*. Preprint, arXiv:1511.02134.

[3] Gmeiner, B., Rüde, U., Stengel, H., Waluga, C., & Wohlmuth, B. (2015). Performance and scalability of hierarchical hybrid multigrid solvers for Stokes systems. *SIAM Journal on Scientific Computing*, *37*(2), C143-C168.

[4] Rudi, J., Malossi, A.C.I., Issac, T., Stadler, G., Gurnis, M., Staar, P.W.J.,… Ghattas, O. (2015). An extreme-scale implicit solver for complex PDEs: highly heterogeneous flow in earth's mantle. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. Austin, TX: ACM.

[5] Stadler, G., Gurnis, M., Burstedde, C., Wilcox, L.C., Alisic, L., & Ghattas, O. (2010). The dynamics of plate tectonics and mantle flow: From local to global scales. *Science*, *329*(5995), 1033-1038.

[6] Sundar, H., Biros, G., Burstedde, C., Rudi, J., Ghattas, O., & Stadler, G. (2012, November). *Parallel geometric-algebraic multigrid on unstructured forests of octrees. Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis* (p. 43). Salt Lake City, UT: IEEE Computer Society Press.

[7] Weismüller, J., Gmeiner, B., Ghelichkhan, S., Huber, M., John, L., Wohlmuth, B.,… Bunge, H.P. (2015). Fast asthenosphere motion in high-resolution global mantle flow models. *Geophysical Research Letters*, *42*(18), 7429-7435.

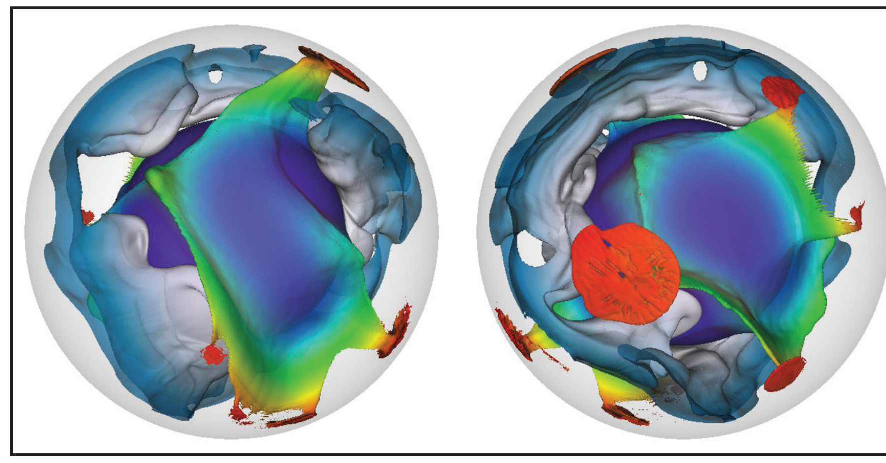


Figure 2. *Two hemispherical views of a global mantle convection model. Cold material (blue to grey colors) sinks towards the core-mantle boundary, while hot low-viscosity material (rainbow colors) rises towards the surface in focused upwellings.*

Wolfgang Bangerth is a professor of mathematics at Texas A&M University. Juliane Dannberg is a graduate student at the GeoForschungsZentrum Potsdam in Germany and a research assistant at Texas A&M University. Rene Gassmöller is a postdoc at both the GeoForschungsZentrum Potsdam and Texas A&M University. Timo Heister is an assistant professor at Clemson University. All four are co-authors of the ASPECT code to simulate mantle convection. Bangerth and Heister are also among the principal authors of the deal.II finite element library.

Compressed Sensing

Continued from page 1

programming problem, for which we have fast algorithms. When the signal is sufficiently sparse, the recovery via ℓ^1 minimization is provably exact [1].

Signals and images are rarely exactly sparse; more often they are only approximately sparse. Additionally, measured data are invariably corrupted by at minimum a small amount of noise; at the very least, small perturbations in the data should cause small perturbations in the reconstruction. So for practical reasons, compressed sensing must be able to deal with both nearly sparse signals and with noise. As was shown by Candès, Romberg, and Tao [4], if x is not exactly sparse, then the quality of the recovered signal is as good as if one knew the location of the largest values of x ahead of time and decided to measure those directly. In other words, the reconstruction is nearly as good as that provided by an oracle which, with full and perfect knowledge about x, extracts the most significant pieces of information.

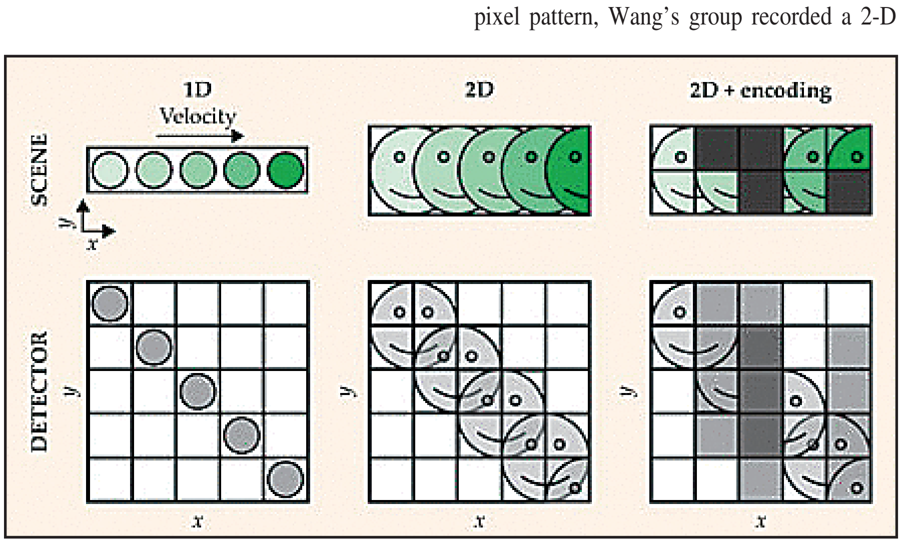


Figure 2. *A one-dimensional scene is sheared vertically and the scene's time domain maps to the detector's y-domain. A movie can be reconstructed by taking each row of pixels as a frame.*

The 2-D Camera

The 2-D camera, which can record non-repetitive, time-evolving events at one hundred billion frames per second, was developed by Wang and his colleagues using a dynamic 2-D imaging technique called *compressed ultrafast photography* (CUP). CUP combines *streak photography*, an ultrafast imaging technique, with compressed sensing [6].

While traditional cameras capture discrete images by repeated opening and closing of a shutter, streak cameras deflect incoming light in a way that maps time onto a spatial dimension on the camera's two-dimensional pixel array. The light is deflected by an amount proportional to its time of arrival, a process known as *shearing*.

Traditional streak cameras can film in only one spatial dimension. When a one-dimensional scene like the one shown in Figure 2—where the object moves from left to right—is sheared vertically, the scene's time domain is translated to the detector's *y*-domain. A movie can then be reconstructed by taking each row of pixels as a frame.

By masking the 2-D scene with a known pixel pattern, Wang's group recorded a 2-D scene in a single shot. The Washington University team was able to achieve this by taking their cues from signal processing. CUP facilitates detection of information about all three dimensions—spatial coordinates x and y, and time t—on a single pixel array by encoding the input scene in the spatial domain. The resultant image is then sheared in the streak camera. A 2-D detector array with a single snapshot measures this encoded and sheared three-dimensional scene (x, y, t). The encoded data can retrieve time information from the input scene in the subsequent image reconstruction.

Applications

The single-shot ultrafast camera is like a microscope for time, one that has many potential applications in fundamental and applied science. One remarkable medical application for the camera is in magnetic resonance imaging (MRI). MRI recovers images from a human body by taking a large number of measurements before reprocessing the data. The scan can take extended periods of time, much to the disadvantage of patients who have to be kept completely still. In the worst cases, the individual's breathing is stopped, depriving the patient of oxygen for the entire scan. Compressed sensing techniques alleviate this issue by reconstructing the image using many fewer samples. These techniques significantly reduce the number of measurements required for the MRI from minutes to mere seconds.

CUP can potentially be coupled with microscopes to enable the observation of transient events in cell structures. For instance, it could be used to observe energy metabolism in cellular mitochondria. It could also enable understanding of light's passage through tissues, which could yield important insights into therapies that use lasers to destroy abnormal or diseased tissue, with the objective of keeping healthy tissue unharmed.

As Wang puts it, the generic nature of the camera means it can be coupled with a wide variety of tools—from microscopes to telescopes—and thus be used to film anything from cellular functions to collapsing supernovae [7].

Acknowledgments

This article is based on a presentation by Dr. Fleming Crim, NSF Assistant Director for Mathematical and Physical Sciences, for the MPS Advisory Committee on November 5, 2015. The author acknowledges input from Dr. Tie Luo, Acting Deputy Division Director, NSF/MPS/Division of Mathematical Sciences.

References

[1] Candès, E., & Romberg, J. (2007). Sparsity and incoherence in compressive sampling. *Inverse Prob., 23*(3), 969985.

[2] Candès, E., & Tao, T. (2006, December). Near optimal signal recovery from random projections: Universal encoding strategies? *IEEE Trans. Inform. Theory, 52*(12), 5406–5425.

[3] Candès, E., Romberg, J., & Tao, T. (2006, February). Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory, 52*(2), 489–509.

[4] Candès, E, Romberg, J., & Tao, T. (2006, August). Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math., 59*(8), 1207–1223.

[5] Donoho, D. (2006, April). Compressed sensing. *IEEE Trans. Inform. Theory, 52*(4), 1289–1306.

[6] Gao, L., Liang, J., Li, C., & Wang, L.V. (2014, December 4). Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature, 516,* 74–77. doi:10.1038/nature14005.

[7] Smart, A.G. (2015, February). An ultrafast camera films light at light speed. *Physics Today, 68*(2), 12. doi: http://dx.doi.org/10.1063/PT.3.2673.

Hans Kaper, founding chair of SIAG/MPE and editor-in-chief of SIAM News, *is an adjunct professor of mathematics at Georgetown University.*

Hot Hands, Streaks and Coin-flips: How *The New York Times* Got it Wrong

By Dan Gusfield

The existence of "Hot Hands" and "Streaks" in sports and gambling is hotly debated. Two recent articles in *The New York Times* (*NYT*) discussing streaks and hot hands in basketball and coin flips misinterpreted elementary concepts in probability and statistics. While it is disheartening to see the newspaper of record make such basic errors in mathematical reporting, the articles do provide a case study for how they got it wrong. This article is adapted from a longer piece I wrote for that purpose and audience [2].

The starting point is an article by George Johnson in the *NYT Sunday Review* on October 18, 2015, titled "Gamblers, Scientists and the Mysterious Hot Hand,"[1] which discusses recent claims by Joshua Miller and Adam Sanjurjo [3] that a classic 1985 paper [1] debunking the concept of hot hands in basketball, "is flawed by a subtle misperception about randomness." Then, on October 27, 2015, in a follow-up *NYT* article (published in *The Upshot*) called "Streaks Like Daniel Murphy's Aren't Necessarily Random,"[2] Binyamin Appelbaum writes that Miller and Sanjurjo [3] claimed that the classic paper had "made a basic statistical error."

What I discuss here is not hot hands per se, but how the *NYT* articles addressed probability and statistics. The following quotes are from the two articles:

(From Johnson): "For a 50 percent shooter, for example, the odds of making a basket are supposed to be no better after a hit – still 50-50. *But in a purely random situation, according to the new analysis, a hit would be expected to be followed by another hit less than half the time.*" (Italics added). The *NYT* article concerns a "purely random situation," not some basketball-related phenomenon. I interpret the Johnson statement in italics as "... the probability that a hit will be followed by another hit is less than one-half."

(From Appelbaum): "Flip a coin, and there's an equal chance it will land heads or tails ... But Joshua Miller of Bocconi University and Adam Sanjurjo of the Universidad de Alicante pointed out something surprising: *In the average series of four coin flips, the sequence heads-heads is significantly less common than heads-tails.* On average, just 40.5 percent of the heads are followed by another heads." (Italics added).

In a "purely random situation," *every* flip will be an H with the same probability that it is a T – exactly one-half. Thus, a hit is *expected* to be followed by another hit (H) one-half of the time, which is as often as it is expected to be followed by a miss (T). So what is going on?

Following a similar example and table in the Miller and Sanjurjo paper (but not yielding a similar conclusion), Johnson did the following. He looked at the 16 length-four sequences shown in Figure 1. For each sequence that contains an H in one of the first three positions (there are 14), he calculated the *percentage* of those Hs that are followed by another H. Call that the *HH-percentage*. Then he calculated the *unweighted* average of the HH-percentages and got about 40.5%.

The arithmetic is right, but so what? The unweighted average gives *equal* weight to each sequence, ignoring the fact that some sequences have more Hs than others, and that occurrences of HH are not uniformly distributed among the sequences. To correctly calculate the probability that an H follows an H, you need to give equal weight to each H that occurs in the first three positions of a sequence. Equivalently, if you start from the HH-percentages of the 14 sequences, you need to compute a *weighted* average of those HH-percentages: each HH-percentage multiplied by the number of Hs in the first three positions of the associated sequence. The *NYT* articles get this wrong, because they suggest that the *unweighted* average of the HH-percentages *equals* the probability of an H following an H in a fair coin flip.

While Johnson and Appelbaum miss the issue of weighted versus unweighted averaging, Miller and Sanjurjo (MS) do not. They state that "The key ... is that it is not the flip that is treated as the *unit of analysis*, but rather the *sequence* of [four] flips from each coin ... Therefore, in treating the sequence as the unit of analysis, the average empirical probability across coins amounts to an *unweighted* average."

The 16 sequences	Number of Hs in the first three positions	Number of HHs	HH-Percentage
HHHH	3	3	100
HHHT	3	2	66.66 ...
HHTH	2	1	50
HHTT	2	1	50
HTHH	2	1	50
HTHT	2	0	0
HTTH	1	0	0
HTTT	1	0	0
THHH	2	2	100
THHT	2	1	50
THTH	1	0	0
THTT	1	0	0
TTHH	1	1	100
TTHT	1	0	0
TTTH	0	0	0
TTTT	0	0	0
Total	24	12	
Average from the first 14 sequences			40.5

Figure 1. *The 16 sequences of length four. The number of Hs in the first three positions is 24, and the number of those Hs followed by another H is 12, exactly 50%. However, the HHs are not distributed uniformly. For example, sequence HHHH has three HHs and TTHH only has one, but both have HH-percentage of 100%. So, the unweighted average of the HH-percentages is not 50%, but about 40.5%. True, but so what? It does not follow that the probability of a hit following another hit is less than half.*

But Why?

Why did MS intentionally calculate *unweighted* averages? Let me explain one reason. Suppose that a player, with an established 50% hit rate, shoots four times in a game. To decide if he had a hot hand in the game, we compute his HH-percentage for the game and compare it to a reference number representing the *expected* HH-percentage for someone *without* a hot hand. We could compare to 50%, which is essentially (but not exactly) what the authors did in the classic 1985 paper. But the MS approach is to model a player *without* a hot hand as a fair coin; we can think of that player as *selecting* (with equal probability) one of the 16 four-flip sequences.[3] The expected value of those 14 HH-percentages is their unweighted average (about 40.5%).

In general, for any k, the unweighted average HH-percentage over the 2^k sequences of length k is less than one-half. So, in the MS view, an HH-percentage of 50% in a game for a player with a hit-rate of 50% is evidence the player *did* exhibit a hot hand, rather than evidence against it. That is the main point made by MS, and the reason for their use of the unweighted average. The *NYT* articles missed that point.

The Wall Street Journal also addressed the hot hands dispute in "The 'hot hand' debate gets flipped on its head,"[4] by Ben Cohen, September 29, 2015, and initially made the same mistake as the *NYT* articles.

1 http://www.nytimes.com/2015/10/18/sunday-review/gamblers-scientists-and-the-mysterious-hot-hand.html

2 http://www.nytimes.com/2015/10/27/upshot/trust-your-eyes-a-hot-streak-is-not-a-myth.html

3 In MS they restrict to the 14 sequences that have an H in one of the first three positions.

4 http://www.wsj.com/articles/the-hot-hand-debate-gets-flipped-on-its-head-1443465711

See **Hot Hands** *on page 5*

Exploration Seismology

Continued from page 2

Another line of research focuses on relaxing the physics and putting more emphasis on fitting the data. The conventional approach insists on obeying the physics for a given set of parameters by solving the PDE $\mathcal{L}(m)u_i = q_i$ and finding the parameters such that u_i fits the data. Instead, we can relax the constraints and construct solutions u_i that fit the data but fail to obey the physics, i.e., $\mathcal{L}(m)u_i \approx q_i$. The goal is now to find parameters m so the solution u_i obeys the physics. We state the problem here as follows: *For given measurements* d_i*, determine the coefficients m such that* $\mathcal{P}u_i \approx d_i$ *and* $\mathcal{L}(m)u_i \approx q_i$ *for* $i = 1,2,\ldots k$.

Such approaches, which place the data and physics on equal footing, are well-known in data assimilation but new in inverse problems. They can be used to solve the original inverse problem while being less sensitive to the initial guess.

Dimensionality Reduction

We can formulate a *multi-experiment* inverse problem generically as

$$\min_m \frac{1}{k}\sum_{i=1}^{k} f_i(m),$$

where f_i measures the data fit for given parameters m. Evaluation of a single f_i requires the solution of a PDE which constitutes the dominant computational cost when solving this optimization problem. The idea is to replace the objective by an unbiased approximation

$$\frac{1}{k}\sum_{i=1}^{k} f_i(m) \approx \frac{1}{|I|}\sum_{i\in I} f_i(m),$$

where $I \subset \{1,2,\ldots,k\}$ is a randomly-chosen subset of size $|I| \ll k$. Using a relatively small number of terms can obtain very good results and lead to an order of magnitude speedup. To guarantee convergence to a solution of the full problem, special optimization techniques have to be developed. Of special interest are techniques to adaptively choose the number of samples based on the required accuracy [2].

Estimating Nuisance Parameters

Many formulations of the inverse problem involve additional *nuisance* parameters that may not be of primary interest but are crucial for finding a meaningful reconstruction. Such parameters include calibration weights or characteristics of noise, such as variance. Solving for these additional parameters alongside the primary ones leads to a *bi-level* optimization problem

$$\min_{m,w} f(m,w).$$

Rather than solve this as a generic non-linear optimization problem, a more attractive approach involves introducing an optimal value function $\bar{f}(m) = \min_w f(m,w)$ and solving a reduced problem for m alone. In many instances optimization in w is easy, and it turns out that the derivatives of $\bar{f}$ with respect to m do not involve derivatives of w with respect to m. In particular, $\partial_m \bar{f}(m) = \partial_m f(m,\bar{w})$ where $\bar{w}$ is the optimal w, implicitly defined through $\partial_w f(m,w) = 0$. Employing the chain rule easily verifies the latter statement, but similar statements can be made when f is not smooth in w (under suitable assumptions). This results in an extremely powerful framework for handling additional parameters in the context of large-scale optimization.

The aforementioned challenges of the seismic inverse problem call for unconventional reformulations of the inverse problem and new computational techniques to handle the large amounts of data. While some of the challenges are unique to exploration seismology, other issues are generally encountered in inverse problems with wave equations. Being able to quantify the uncertainty in the solution is important in many applications. Together with faster methods to solve the inverse problem, this may ultimately lead to computationally-feasible approaches for uncertainty mitigation.

References

[1] Aravkin, A., & Van Leeuwen, T. (2012). Estimating nuisance parameters in inverse problems. *Inverse Problems, 28*(11), 115016.

[2] Aravkin, A., Friedlander, M.P., Herrmann, F.J., & Van Leeuwen, T. (2012). Robust inversion, dimensionality reduction, and randomized sampling. *Mathematical Programming, 134*(1), 101–125.

[3] Martin, J., Wilcox, L.C., Burstedde, C., & Ghattas, O. (2012). A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion. *SIAM Journal on Scientific Computing, 34*(3), A1460–A1487.

[4] Symes, W.W. (2014). Seismic inverse problems: recent developments in theory and practice. In A.K. Louis, S. Arridge, & B. Rundell (Eds.) *Proceedings of the Inverse Problems from Theory to Applications Conference* (pp. 2-6). Bristol, UK: IOP Publishing.

[5] Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics, 49*(8), 1259–1266.

[6] Van Leeuwen, T., & Herrmann, F.J. (2014). 3D Frequency-Domain Seismic Inversion with Controlled Sloppiness. *SIAM Journal on Scientific Computing, 36*(5), S192–S217.

[7] Van Leeuwen, T., & Herrmann, F.J. (2016, January). A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems, 32*(1), 015007.

Tristan van Leeuwen is an assistant professor in the Department of Mathematics at Utrecht University, where he researches inverse problems in imaging. In July 2015, Van Leeuwen received the SIAM Geosciences Junior Geoscientist prize for his research on computational methods for seismic inversion.

Julia: A Fast Language for Numerical Computing

By Alan Edelman

The Environmental Protection Agency's discovery that Volkswagen installed emissions-cheating software in over 11 million vehicles last September garnered much attention. According to reports, the software made it seem like the vehicles were releasing significantly less nitrogen oxides than in reality, allowing them to discharge up to 40 times the allowable amount of smog-producing pollution. This prompted *The New York Times* to call for an openness in software that goes beyond the scientific need for transparency. Software openness is becoming a matter of safety and accountability for compliance purposes. This need applies to more than just cars, as the Internet of Things requires communication of data and instructions among multitudes of devices. Open source is not just a checkbox; easy-to-read common code is increasingly becoming paramount.

With its speed, openness, and transparency, Julia is emerging as that "must-have" easy-to-read common language. And in terms of safety and accountability, Julia is being used as the specification language for the next-generation airplane collision avoiding system[1] (see Figure 1).

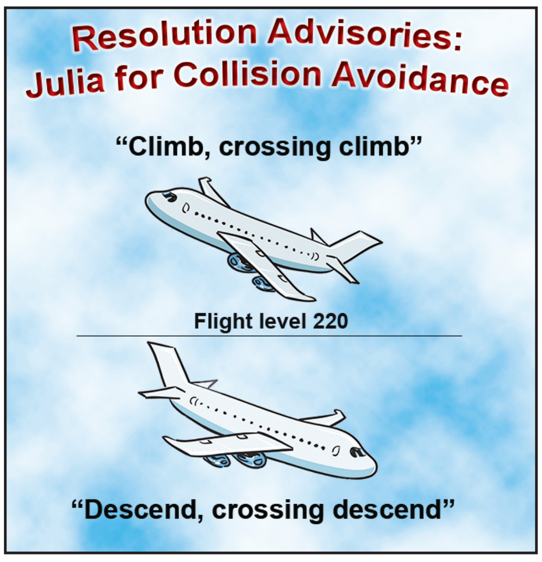


Figure 1. *Julia is being used in the development and specification of the next-generation aircraft collision avoidance system.*

Why use Julia over languages such as R or Python? The differences lie in the programs' intricacies. These are all very good languages, but when programming in R or Python the two-language problem lurks: the need to prototype with one slow dynamic language and rewrite with a fast static language to obtain the final product. Having been around longer than Julia, R and Python have naturally established the kind of popularity that takes some years to acquire. Nevertheless, the two-language problem is very serious. Prototyping in R or Python necessitates a rewrite for speed and deployment. If emissions software is written in Python, one must quickly drill down to C code to understand the details of what is happening. Julia's advantage is that it solves the two-language problem.

While exact usage statistics are difficult to generate, Julia's popularity is doubling every nine months, Moore's law style. There are currently an estimated 100,000 Julia users, nearly 1,000 Julia packages, and more than 400 GitHub contributors. Dozens of universities are also using Julia in classes in engineering, big data, optimization, and numerical analysis, as well as in undergraduate computer science, linear algebra, and statistics. Anecdotally, many engineering and financial companies are using Julia at the "grassroots" level.

Nick Trefethen, former SIAM President and my Ph.D. advisor, first got me thinking about how to best articulate Julia's speed and advantages. I spoke about Julia at the "New Directions in Numerical Computation" linear algebra conference in honor of Trefethen's 60th birthday this past August. The following commentary includes a few key points from that talk.

Placing Emphasis on Performance

Julia demonstrates that if designed correctly, the very features that help the computer can also appeal to the human. Those unfamiliar with the transformations of code tend to replace complete understandings with magic accelerators based on some partial truths. Users hear that compiling makes code run faster. Jitting makes code run faster. Vectorizing makes code run faster. Declaring types makes code run faster.

Interestingly, what makes an algorithm run faster seems less mysterious than what makes code run fast. Algorithms run quickly when they do not waste time computing unnecessary things. This efficiency can mean taking advantage of structure (as in not multiplying by or adding zeros in sparse matrix computations), or computing only as many digits as are necessary (in iterative methods, or by truncating to take advantage of hardware).

Ultimately, code runs faster when it is not beleaguered by unnecessary computational baggage. That the computer language is thought of separately from the computer algorithm is unfortunate. When considering numerical error, every computation must work just right. If not, the entire computation goes wrong. The same can be said of performance; a language must be designed just right for speed.

The earliest computer languages were designed primarily to meet the needs of the computer, and were close enough to the hardware to be relatively fast. Then came interpreted languages, which were designed primarily to meet the needs of the human. This facilitated productivity and was thus the better choice for most people. Julia finds a new place by striking a reasonable balance between human and computer.

```
# bat functions: semicircle, ellipse, shoulders, bottom, cowl
σ(x) = √(1-x.^2)
e(x) = 3σ(x/7)
s(x) = 4.2 - .5x - 2.8σ(.5x-.5)
b(x) = σ(abs(2-x)-1)-x.^2/11 + .5x - 3
c(x) = [1.7, 1.7, 2.6, .9];

# plot symmetrically across y-axis (NaN's lift the "pen")
using PyPlot
p(x,f) = plot([-x,0,x],[f(x),NaN,f(x)])
p(3:.1:7,e);p(4:.1:7,t->-e(t))         # ellipse
p(1:.1:3,s);p(0:.1:4,b)                # shoulders and bottom
p([0,.5,.8,1],c)                       # cowl is piecewise linear
```

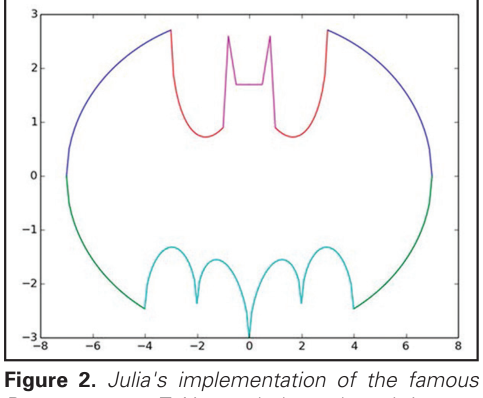


Figure 2. *Julia's implementation of the famous Batman curve. TeX symbols such as \sigma + <tab> display beautifully.*

People initially thought, almost by some kind of conservation law of nature, that a dynamic mathematical language was required to be slow. Julia shows that a carefully designed language that is comfortable for the programmer can still run quickly. It is important to understand that to date, many interpreted computer languages were designed with use cases in mind, while performance and acceleration came grafted in as potential "add-ons." Placing more emphasis on performance enhances the cooperation between human and computer.

[1] Robert Moss: Julia and the Next Generation Airborne Collision Avoidance System: https://www.youtube.com/watch?v=19zm1Fn0S9M

A First Julia Experience: the Batman Curve

The famous Batman curve is pictured in Figure 2, which users can experiment with on juliabox.org. The expressions in Figure 2 are equivalent to others that may be found online; I believe this version is much easier for people to read (See interactive code in Lecture 1).[2]

Julia is Fast and Flexible

Computer programs run faster when they transform to a tight set consisting of a small number of quick execution steps. Consider some uses of the plus character "+." We ask so very much from plus, and know which plus we mean by context. Assembly language may use different instructions to express the various additions on a machine. The good news is that aside from some pushing, popping, and moving, one can see how tight Julia's instructions are (see Figure 3).

Of course, "+" can refer to so much more, including dense matrix addition, sparse matrix addition, structured matrix addition, and any type of combination. The point is that Julia produces tight code in all instances.

How does Julia do it? For those desiring a quick answer, "multiple dispatch" is the explanation. The Julia language is built on careful principles of type stability that allow for type inference, which enables Julia to perform proper code selection through the multiple dispatch mechanism. Julia reduces the uncertainty in a computation that can waste time.

One can see multiple dispatch in action with the "@which" command in Figure 4, which allows users to drill into Julia code if they wish to see how it implements various "+" symbols. Without the user declaring types, the Julia program knows which "+" is intended and uses the correct assembly code without wasting time.

```
In [1]: x=1.0; y=1.0;
        x+y
        @code_native(x+y)

                .text
        Filename: float.jl
        Source line: 120
                push    RBP
                mov     RBP, RSP
        Source line: 120
                vaddsd  XMM0, XMM0, XMM1
                pop     RBP
                ret

In [2]: x=1;y=1;
        x+y
        @code_native(x+y)

                .text
        Filename: int.jl
        Source line: 33
                push    RBP
                mov     RBP, RSP
        Source line: 33
                add     RDI, RSI
                mov     RAX, RDI
                pop     RBP
                ret
```

Figure 3. *Julia users rarely look at assembly, but they readily can. Julia's tight assembly is a clue as to why it is so fast.*

[2] Use Numerical Computing with Julia: https://github.com/alanedelman/18.337_2015

```
In [1]:  @which(1+1)
Out[1]:  +(x::Int64,y::Int64) at int.jl:33

In [2]:  @which (1.0+1.0)
Out[2]:  +(x::Float64,y::Float64) at float.jl:120

In [3]:  @which(1+1.0)
Out[3]:  +(x::Number,y::Number) at promotion.jl:158

In [4]:  @which(rand(10)+1)
Out[4]:  +(A::AbstractArray{T,N},x::Number) at array.jl:774

In [5]:  @which(1//2+1//3)
Out[5]:  +(x::Rational{T<:Integer},y::Rational{T<:Integer}) at rational.jl:113
```

Figure 4. *Multiple dispatch in action with the "@which" command.*

Julia is very much Julia all the way down. (There are bits of other software, and LLVM at the very bottom.) While Julia uses and plays nicely with packages from Python and other computing languages, it does not need to for speed. Even packages such as LAPACK would likely run as fast and could be more flexible if rewritten in Julia. Other numerical languages are front ends to C, C++, or Fortran.

Open Source Research

In reality, Julia is fast because the community cares about performance. Because it is an open project, self-described "performance obsessive types" have flocked to the project, thus boosting Julia's performance. I personally want to see Julia emerge as the language of high performance computing, blurring the distinction between "Silicon Valley"-style big data and artificial intelligence (AI) computations, Wall Street style number crunching, and National Laboratory style scientific computing.

Numerical software research enjoys the benefit of its fruits being quite practical. Those with decades of experience in the business know that users have been sharing code development long before GitHub or the use of the term "open source."

It has been a pleasure to nurture the Massachusetts Institute of Technology's component of this worldwide effort and to experience the brilliant contributions from so many first-class researchers. Experience Julia for yourself by searching "Juliacon2015" online and watching the videos, trying juliabox.org, or visiting julialang.org.

Acknowledgments: I'd like to dedicate this article to my Ph.D. adviser and former SIAM President, Nick Trefethen, on the occasion of his 60th birthday.

Alan Edelman is a professor of applied mathematics at Massachusetts Institute of Technology. He has been working in the area of high performance computing systems, networks, software, and algorithms for 30 years and has won many prizes for his work. With Julia he sees a fresh approach to high performance technical computing.

Hot Hands

Continued from page 4

Cohen wrote: "Toss a coin four times. Write down what happened. Repeat that process one million times. What percentage of flips after heads also come up heads? The obvious answer is 50%. That answer is also wrong. The real answer is 40%."

But on September 30, in an online version of the article, the error was corrected to "Toss a coin four times. Write down the percentage of heads on the flips coming immediately after heads. Repeat that process one million times. On average, what is that percentage?"

The *NYT*, on the other hand, has not yet issued a correction at the time of this publication. As an educator in a field involving mathematical reasoning, and one concerned with the public's understanding of quantitative issues, this is disturbing. Articles such as this reinforce the need for discussions in high school and college focused on quantitative reasoning, data analysis, probability, and statistics.

References

[1] Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295–314.

[2] Gusfield, D. (2015, December 29). Hot hands, streaks and coin-flips: Numerical nonsense in the New York Times. *Cornell University Library*. Preprint, arXiv:1512.08773v1 [math.HO].

[3] Miller, J.B., & Sanjurjo, A. (2015, September 15). Surprised by the Gambler's and Hot Hand Fallacies? A Truth in the Law of Small Numbers. *IGIER Working Paper no. 552.*

Dan Gusfield is a professor in the computer science department at the University of California, Davis.

Addressing Challenges in Reduced-Order Modeling

By Kevin Carlberg

One of applied mathematics' great contributions is the foundation it provides for simulating physical phenomena. From the derivation of consistent, stable, and convergent discretization schemes to the development of efficient parallel solvers, mathematical advances have enabled the ubiquitous nature of modeling and simulation in applications ranging from protein-structure prediction to aircraft design. Today, the predictive capability of validated computational models allows simulation to replace physical experimentation in many scenarios, which facilitates the realization of deeper analyses and better designs at lower costs. However, there is a catch: the resolution required to achieve such high fidelity leads to large-scale models whose simulations can consume weeks on a supercomputer. This creates a massive gap between the simulation times of high-fidelity models and the rapid time-to-solution demands of *time-critical* (e.g., real-time analysis) and *many-query* (e.g., uncertainty quantification) applications in engineering and science.

To bridge this gap, researchers have pursued *reduced-order modeling*—which integrates techniques from data science, modeling, and simulation—as a strategy for reducing the computational cost of such models while preserving high levels of fidelity. First, these methods execute analyses (e.g., simulating the model, solving Lyapunov equations) during an offline 'training' stage; these analyses generate data that can be mined to extract important physical features, such as low-dimensional solution manifolds and interpolation points for approximating nonlinear functions. Next, these techniques reduce the dimensionality and computational complexity of the high-fidelity model by projecting the governing equations onto the low-dimensional manifold and introducing other approximations where necessary. The resulting *reduced-order model* (ROM) can then be rapidly simulated during an online 'deployed' stage.

While significant advances have been made in reduced-order modeling over the past fifteen years, many outstanding challenges face the community, especially with respect to applying model reduction to parameterized nonlinear dynamical systems. The West Coast ROM Workshop—held last November at Sandia National Laboratories in Livermore, California—brought together local researchers from both academia and the national laboratories to address these challenges. Speakers provided a range of interesting perspectives on the topic.

One workshop theme focused on applying ROMs to truly large-scale nonlinear problems in engineering and science. To motivate this, invited speaker Charbel Farhat provided a number of compelling examples in which the computational cost incurred by such models poses a major bottleneck to design engineers across the naval, aerospace, and automotive industries. A number of challenges arise in this case. First, ROM techniques must be tightly integrated with the original high-fidelity simulation code because most nonlinear ROM methods realize computational savings by performing computations with the high-fidelity model on a small subset of the computational domain. Second, ensuring accurate ROM solutions can be challenging due to the complex dynamics (e.g., stiffness, chaoticity) exhibited by many large-scale dynamical systems. Finally, when the model is very large scale, the computational costs of both the offline training and online deployment can remain prohibitive; devising ways to reduce them is often essential.

Farhat presented promising results for the energy-conserving sampling and weighting (ECSW) ROM method [5] applied to large-scale problems in structural dynamics (Figure 1). This method integrates with finite-element codes by performing potential-energy computations on a subset of the domain's elements; it produces accurate solutions by ensuring the ROM inherits the energy-conservation principle of the high-fidelity model. Jeffrey Fike presented results on integrating nonlinear ROM methods with a compressible fluid-dynamics code (developed at Sandia), and Irina Tezaur proposed a supporting approach for improving the accuracy of compressible-flow ROMs by rotating typical solution subspaces to include modes needed for energy dissipation [1]. Kyle Washabaugh presented an approach for robustly deforming a

See **Reduced-Order Modeling** *on page 8*

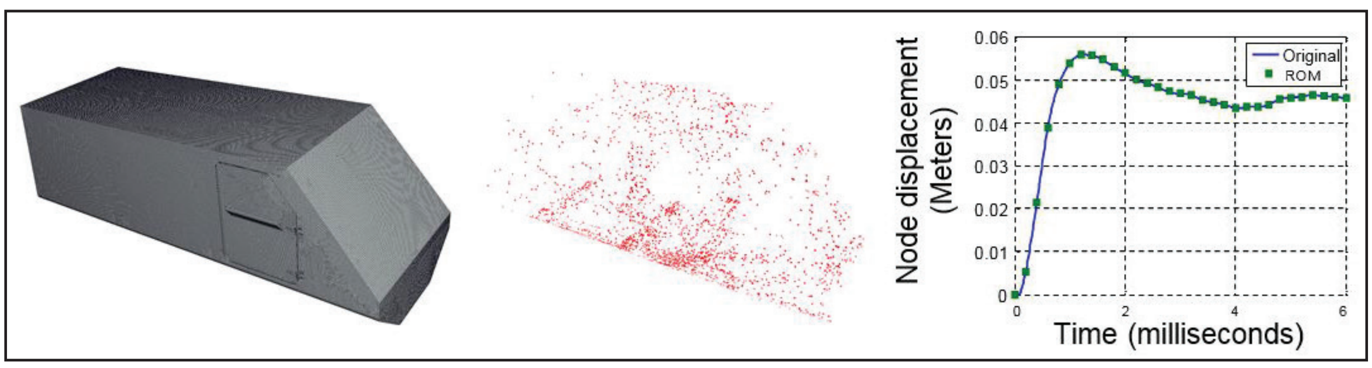


Figure 1. *Simulating the underbody blast of a V-hull vehicle with a ROM. Original computational domain with 2.4 × 10⁵ elements (left), subset of the domain with 2 × 10³ elements used by the ROM (center), and a comparison of the results (right). The ROM generated sub-1% displacement errors with a 10⁴ wall-time speedup. Courtesy C. Farhat.*

Will AI Make Jobs Obsolete?

BOOK REVIEW
By James Case

Rise of the Robots: Technology and the Threat of a Jobless Future. *By Martin Ford, Basic Books, New York, 2015, 354 pages, $28.99.*

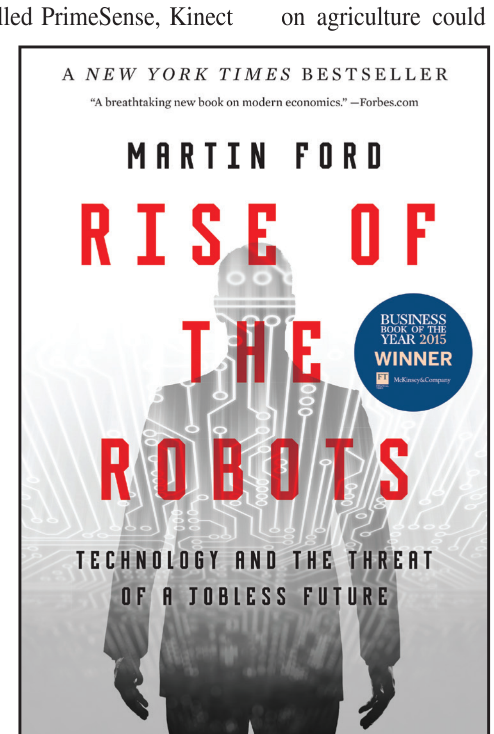
Martin Ford's recent book, *Rise of the Robots: Technology and the Threat of a Jobless Future*, explores the impact of automation, artificial intelligence, and smart machines on our economy and society. Ford is the founder of a Silicon Valley-based software development firm, with twenty-five years of experience in computer design and software development. The book's early chapters describe recent developments in information technology, with emphasis on the synergy between hardware improvement and algorithm design. Later chapters explore the implications for specific businesses, industries, and sectors of society.

Chapter 1 focuses on recent advances in computer vision and the shop floor revolution they portend. A firm called Industrial Perception, Inc. is currently marketing a robot capable of restacking an irregular heap of boxes of varying size, shape, color, and orientation. A human assigned such a task would immediately know where to begin. But robots, lacking applicable experience, must pause to formulate a plan. As a result, writes Ford, Industrial Perception's machine appears sluggish and hesitant at times. However, the company estimates that upgrades already in the pipeline will permit the transfer of a box per second, far faster than the six seconds apiece required by human workers.

The electric car company Tesla uses 160 surprisingly versatile robots to assemble some 400 cars per week in its Fremont, California factory. Each robot is able to exchange one tool for another before starting a new task. The same robot that installs the seats, for instance, is able to retool before applying adhesive and dropping the windshield into place.

Operative, vision-equipped robots, including those used by Tesla, currently see in just two dimensions. But modern technological developments indicate the possibility for more advanced vision. In November 2006, Nintendo introduced its Wii video game console, complete with a wireless wand enabling players to exert control by waving the wand to return an approaching tennis ball, for instance. Sony soon entered the market with a competing product, likewise reliant on two-dimensional vision. Microsoft, on the other hand, gambled with an add-on to the Xbox 360 game console called Kinect, which sees in three dimensions.

Using technology developed by a small Israeli company called PrimeSense, Kinect deploys a sort of sonar that operates at the speed of light. Bathing its field of vision in flashing infrared, Kinect measures the length of time required for each reflected flash to return to (roughly) its point of origin. Players need no magic wand to interact with Kinect; they need only move about in its field of vision. Remarkably, Kinect makes a sophisticated machine vision system—which might previously have cost tens or even hundreds of thousands of dollars—available in a lightweight consumer device priced at $150!



Martin Ford's newest book, Rise of the Robots: Technology and the Threat of a Jobless Future. *Photo credit: Basic Books.*

Researchers in robotics were quick to realize the potential for such technology to revolutionize their field. Within weeks of Kinect's introduction, says Ford, a number of potential competitors had hacked into the system and posted videos on YouTube demonstrating the ability of their own retro-engineered devices to see in three dimensions.

Time alone will reveal the practical impact of acute machine vision, which includes the ability to see in three dimensions. But Parkdale Mills, a textile factory in Gaffney, South Carolina, may offer a glimpse into the future. Employing a mere 140 workers, Parkdale now ships quantities that would have required 2,000 employees as recently as 1980. The impact of machine vision on agriculture could be equally dramatic. Ford points out that experimental robots are already pruning grapevines in France using machine vision augmented by algorithms that decide which branches to eliminate. And a Japanese machine is able to recognize ripe strawberries by color and pick one every eight seconds, working continuously even after dark!

In Chapter 3, Ford cites a study by Martin Grötschel of the Zuse Institute Berlin. The study asserts that a certain complex production planning problem would have taken 82 years to solve in 1982, using what was then state-of-the-art hardware and software. As of 2003, the same problem could be solved in about a minute, an improvement of roughly 43 million! And since computer speed increased "only" a thousand-fold during that time, Ford writes that the remaining factor of 43,000 must be due to algorithm improvement. This suggests that improved algorithm design may pose an even greater threat to mankind's long-term job prospects than do faster computers.

Chapter 4 reproduces a three-paragraph newspaper account of a postseason baseball game played in October 2009. Though not the most eloquent prose, it is eminently readable, grammatically correct, and an accurate description of the event in question. The account is remarkable only because the author is a computer program! The software responsible, called Stats Monkey, was created by a team of students and professionals at Northwestern University's Intelligent Information Laboratory. Its purpose is to automate the process of sports reporting by transforming a stream of numerical data into a compelling narrative, complete with items of human interest.

To realize the commercial potential of their design, the creators of Stats Monkey founded a company called Narrative Science, Inc. They raised venture capital, hired a team of experienced software developers, and set out to construct a far more ambitious system called Quill. A variety of media outlets, including *Forbes*, now employ the system to compose articles on sports, business, and politics. Co-founder Kristian Hammond estimates that 90 percent of all news articles will be written algorithmically within fifteen years. In addition, Ford notes that one of the company's earliest backers was In-Q-Tel, the venture capital arm of the Central Intelligence Agency. Presumably the company's software is already being modified to transform the ever-growing deluge of raw intelligence data into easy-to-digest narrative reports.

The Narrative Science experience illustrates Ford's claim that a wide variety of white-collar jobs face clear and present danger of automation. For further evidence, one

See **Artificial Intelligence** *on page 7*

Artificial Intelligence

Continued from page 6

might invoke Google's introduction of an online language translation tool. The project development team began by assembling millions of pages previously translated into multiple languages, beginning with official documents prepared by the United Nations. Later on they exploited the web, where proprietary search engines located a host of additional pages suitable for analysis by the company's machine-learning algorithms. In 2005, Google's natural language translator was the clear winner of the National Institute of Standards and Technology's annual machine translation contest, besting a host of rival systems constructed by professional linguists. Google's system is not yet competitive with the product of skilled human translators, but it offers two-way translation between upwards of 500 language pairs, meaning that pretty much anyone can freely and all but instantaneously obtain a rough translation of just about any document in thirty or so different languages. While expensive human translators may never become extinct, call for their services seems certain to diminish.

Rise of the Robots is written, as the subtitle indicates, to warn of an impending shortage of jobs. Ford quotes a 2013 study by Carl Frey and Michael Osborne of Oxford University (UK), which found that occupations accounting for nearly half of total U.S. employment may be vulnerable to automation within the next 20 years. He also points to the emergence of companies like YouTube which, at the time of its $1.65 billion acquisition by Google, had been in business for less than two years and employed only sixty-five people. That's $25 million per employee!

In Ford's eyes, the prevalence of soaring per-employee valuations—which apply to a number of other high-tech start-ups as well as YouTube—constitute a direct challenge to the conventional economic wisdom whereby advanced technologies must, in the process of destroying "old-economy" jobs, automatically create "new-economy" substitutes requiring greater skill and offering better pay. He takes little comfort from the fact that things have often worked out that way in the past, since they have not done so lately and probably will not in the future. Indeed, Ford presents mounting evidence that "this time is different," in that the next new economy promises to be of the jobless variety. Accordingly, he believes that nothing short of a radical reorganization of society—what political scientists would call a new "social contract"—will enable the ordinary citizen to earn a living wage.

The sort of reorganization Ford has in mind is outlined in the final chapter, entitled "Toward a New Economic Paradigm." One need not agree with his conclusions or specific reform proposals to be impressed with the wealth of evidence he has marshaled in support of his central thesis. *Rise of the Robots* is well worth reading by anyone involved even peripherally in the information technology revolution, or curious to know where it seems to be leading.

James Case writes from Baltimore, Maryland.

Optics and Photonics Workshop at the SIAM Annual Meeting

By Shari Moskow

We are excited to announce the "NSF-SIAM Optics and Photonics Workshop," which will take place on July 11, 2016, as part of the 2016 SIAM Annual Meeting in Boston, MA. Featuring several accomplished applied mathematicians in the field of optics and photonics, workshop talks will survey the field and its open problems for the broader applied mathematics community. A representative from the NSF Division of Mathematical Sciences will discuss and answer questions about the new NSF program in Optics and Photonics.[1]

Advances in our understanding of the behavior of light are necessary both for improving current technologies and realizing new and envisioned ones. These technologies directly influence a variety of sectors in our society including communications, defense and national security, energy, and medicine. Due to recent developments in computational tools, materials science, and nanostructure fabrication, mathematicians have a larger role than ever to play in this field. This workshop can help applied mathematicians bring their vision of the future of mathematics in optics and photonics to a large mathematical audience. We hope the meeting will encourage collaborations among researchers in this field from a diverse set of scientific backgrounds.

Among others, workshop topics will include the following: (1) Research on light-matter interaction, including—but not limited to—low-loss metamaterials, plasmonics, and quantum phenomena; (2) Multiphysics coupling between classical electromagnetic and quantum mechanical phenomena; (3) The science of light propagation and imaging through scattering, dispersive, and turbulent media, which encompasses advances in radiative transport theory, statistical inverse theory, numerical inversion methods, simulation models, and hybrid imaging models; and (4) Nonlinear photonics and the interplay between nonlinearity and randomness.

These challenges in optics and photonics touch on many areas of mathematics, including partial differential equations, ordinary differential equations, dynamical systems, functional analysis, numerical analysis, geometry, calculus of variations, probability, and stochastic differential equations, among others. Traditionally, models from different optical regimes have been associated with very different mathematical concepts and techniques. The development of the next generation of optical devices will require coupling of these models, and collaborations among mathematicians from various backgrounds.

Thanks to the NSF, we expect to have funding available for researchers from underrepresented groups, graduate students, postdoctoral associates, and junior researchers to travel and participate in the meeting. See the SIAM annual meeting webpage[2] to apply. The workshop will expose students and junior scientists to critical research areas in optics and photonics, as well as relevant and important areas of application of mathematics and computational science.

The workshop is timely in that it precedes the Institute for Mathematics and its Applications' thematic year on Mathematics and Optics.[3] The year-long (2016-2017) program will address the study of optical phenomena and associated areas of applied and computational mathematics, with the goal of connecting mathematical and computational scientists with the interdisciplinary community.

Shari Moskow is a professor and department head of mathematics at Drexel University. She is currently a co-organizer of the IMA thematic year on Mathematics and Optics, and also organized the workshop discussed above.

[1] http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf16004

[2] http://www.siam.org/meetings/an16/workshops.php#optics

[3] www.ima.umn.edu/2016-2017/

Reduced-Order Modeling

Continued from page 6

sample mesh in HPC environments when using ROMs to predict steady-state aerodynamic flows subjected to geometric deformations [11]. In contrast, invited speaker J. Nathan Kutz offered a framework that enables nonlinear model reduction *without* requiring access to the high-fidelity simulation code. The approach, based on Koopman theory, approximates nonlinear dynamical systems with a linear operator constructed from time-history data of dynamical-system *observables* [2]. To reduce offline training costs, Geoffrey Oxberry presented a method that leverages ideas from error control in adaptive time integrators to decrease the amount of data needed to construct low-dimensional solution subspaces [9]. To decrease online simulation times in parallel-computing environments, Kevin Carlberg proposed a new method for time parallelism. This technique applies data-driven ROM-solution forecasts [3] as a coarse propagator in the parareal framework.

A second major workshop theme focused on applying ROMs to design optimization. These many-query problems—which are often formulated as mathematical optimization problems constrained by partial differential equations—can require hundreds of simulations (and sensitivity analyses) of the computational model. Thus, rapid model evaluations are necessary when faced with time or resource constraints. Youngsoo Choi presented one approach for applying ROMs to design optimization [4]. The method adopts the classical offline–online strategy, wherein a database of ROMs for the (linear) model is constructed offline, and these ROMs are interpolated online on appropriate matrix manifolds. This method is amenable to real-time applications because it does not require any online high-fidelity simulations; however, it lacks convergence guarantees and—due to the costly offline stage—requires the number of online optimization iterations to exceed a 'break-even' threshold before computational savings can be realized. Invited speaker Louis Durlofsky proposed a related method based on the trajectory piecewise linear (TPWL) ROM, and showed promising results on oil-production optimization under water injection [6]. Matthew Zahr proposed an alternative approach [12] that eschews the typical offline–online strategy in favor of a trust-region approach, wherein the high-fidelity-model solution (and sensitivities) are computed at trust-region centers (Figure 2). This approach guarantees convergence, but is not amenable to real-time applications due to the 'mixing' of high-fidelity and reduced-order model evaluations during the solution to the optimization problem. Both Zahr and invited speaker Michael Frenklach proposed strategies for reducing the dimensionality of high-dimensional parameter spaces; the former employed an adaptive strategy based on the gradient of the Lagrangian, while the latter applied active subspaces to a combustion-chemistry problem.

Other important topics were addressed, with contributions from Tanya Kostova, who presented a technique employing both the system state and velocity as data in solution-subspace computation [8]; Sumeet Trehan, who has applied statistical learning to construct TPWL error surrogates; Syuzanna Sargsyan, who has developed ROMs that adapt to particular physical regimes [10]; and invited speaker Jaijeet Roychowdhury, who proposed a representation of continuous systems as boolean finite state machines [7].

Despite the many challenges, model reduction remains an exciting research area that is making rapid progress toward bridging the gap between high-fidelity models and time-critical applications in engineering and science.

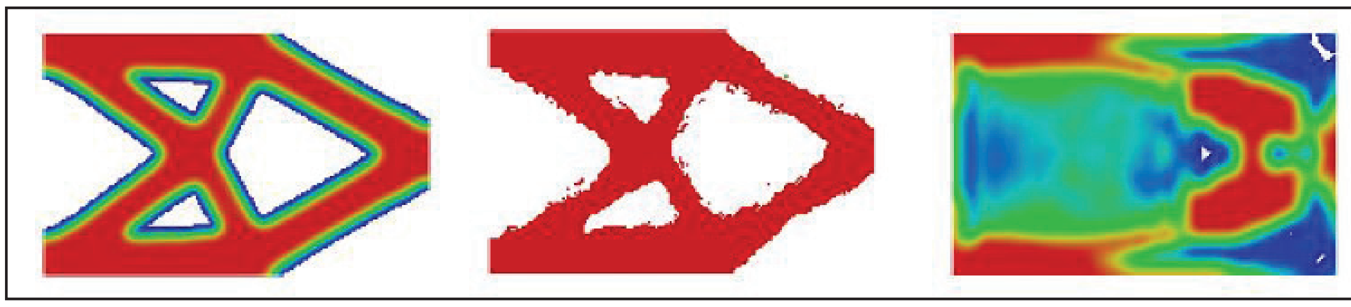


Figure 2. *Topology optimization with reduced-order models. Optimal solution (left), ROM-based optimization solution after 2000 seconds (center), high-fidelity-model-based optimization after 2000 seconds (right). Courtesy M. Zahr.*

References

[1] Balajewicz, M., Tezaur, I., & Dowell, E. (2015). Minimal subspace rotation on the Stiefel manifold for stabilization and enhancement of projection-based reduced order models for the compressible Navier-Stokes equations. *Cornell University Library*. Preprint, arXiv:1504.06661.

[2] Brunton, S.L., Brunton, B.W., Proctor, J.L., & Kutz, J.N. (2015). Koopman observable sub-spaces and finite linear representations of nonlinear dynamical systems for control. *Cornell University Library*. Preprint, arXiv:1510.03007.

[3] Carlberg, K., Ray, J., & Van Bloemen Waanders, B. (2015). Decreasing the temporal complexity for non- linear, implicit reduced-order models by forecasting. *Computer Methods in Applied Mechanics and Engineering*, *289*, 79–103.

[4] Choi, Y., Amsallem, D., & Farhat, C. (2015). Gradient-based constrained optimization using a database of linear reduced-order models. *Cornell University Library*. Preprint, arXiv:1506.07849.

[5] Farhat, C., Avery, P., Chapman, T., & Cortial, J. (2014). Dimensional reduction of nonlinear finite element dynamic models with finite rotations and energy-based mesh sampling and weighting for computational efficiency. *International Journal for Numerical Methods in Engineering*, *98*(9), 625–662.

[6] He, J., & Durlofsky, L.J. (2014). Reduced-order modeling for compositional simulation by use of trajectory piecewise linearization. *SPE Journal*, *19*(05), 858–872.

[7] Karthik, A.V., Ray, S., Nuzzo, P., Mishchenko, A., Brayton, R., & Roychowdhury, J. (2014). ABCD-NL: Approximating continuous non-linear dynamical systems using purely boolean models for analog/mixed- signal verification. *Design Automation Conference (ASP-DAC)*, 2014 19th Asia and South Pacific (pp. 250-255). Singapore: IEEE.

[8] Kostova, T., Oxberry, G., Chand, K., & Arrighi, W. (2015). Error bounds and analysis of proper orthogonal decomposition model reduction methods using snapshots from the solution and the time derivatives. *Cornell University Library*. Preprint, arXiv:1501.02004.

[9] Oxberry, G., Kostova-Vassilevska, T., Arrighi, B., & Chand, K. (2015). *Limited-memory adaptive snapshot selection for proper orthogonal decomposition*. Livermore, CA: Lawrence Livermore National Laboratory (LLNL).

[10] Sargsyan, S., Brunton, S.L., & Kutz, J.N. (2015). Nonlinear model reduction for dynamical systems using sparse sensor locations from learned libraries. *Physical Review E*, *92*(3), 033304.

[11] Washabaugh, K., Zahr, M.J., & Farhat, C. (2016). On the Use of Discrete Nonlinear Reduced-Order Models for the Prediction of Steady-State Flows Past Parametrically Deformed Complex Geometries. *54th AIAA Aerospace Sciences Meeting* (p. 1814). San Diego, CA: AIAA SciTech.

[12] Zahr, M.J., & Farhat, C. (2015). Progressive construction of a parametric reduced-order model for PDE-constrained optimization. *International Journal for Numerical Methods in Engineering*, *102*(5), 1111–1135.

Kevin Carlberg is a Principal Member of Technical Staff at Sandia National Laboratories in Livermore, California. His research focuses on developing model-reduction techniques for large-scale nonlinear dynamical systems.

Classroom Notes: Symmetric Matrices and (a Little) Work

MATHEMATICAL CURIOSITIES

By Mark Levi

Here are a few unpretentious observations that occurred to me several years ago after teaching a linear algebra course. I am not making any claim to their originality.

1. The elegant but also a bit antiseptic definition of a symmetric $n \times n$ real matrix A as the one satisfying the identity

$$(A\mathbf{x}, \mathbf{y}) - (\mathbf{x}, A\mathbf{y}) = 0 \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \tag{1}$$

has a physical interpretation: this identity is equivalent to saying that the work done by the linear force field $\mathbf{F}(\mathbf{x}) = A\mathbf{x}$ around the parallelogram generated by $\mathbf{x}$ and $\mathbf{y}$ vanishes (see Figure 1).

Figure 2 illustrates proof of this equivalence. The average forces on each of the sides of the parallelogram are equal to the forces $\mathbf{F}_i$ at the midpoints M_i. The total work W around the parallelogram, grouping parallel sides together, is

$$W = (\mathbf{F}_1 - \mathbf{F}_3, \mathbf{x}) + (\mathbf{F}_2 - \mathbf{F}_4, \mathbf{y});$$

and since $\mathbf{F}_1 - \mathbf{F}_3 = -A\mathbf{y}$ and $\mathbf{F}_2 - \mathbf{F}_4 = A\mathbf{x}$,

this gives $W = (A\mathbf{x}, \mathbf{y}) - (\mathbf{x}, A\mathbf{y})$.

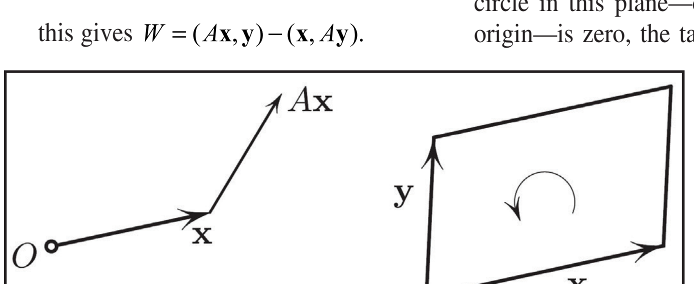


Figure 1. *Vector field A**x** and the closed parallelogram path.*

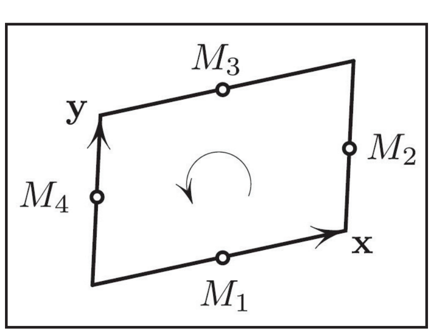


Figure 2. *Physical meaning of (A**x**, **y**) – (**x**, A**y**).*

In particular, (1) expresses the conservativeness of the vector field $A\mathbf{x}$.

2. Here is a physical reason why eigenvalues of a symmetric matrix are real. Assuming for a moment that they are not, consider the plane spanned by the real and the imaginary parts $\mathbf{u}$, $\mathbf{v}$ of the eigenvector $\mathbf{w} = \mathbf{u} + i\mathbf{v}$. At each point $\mathbf{x}$ in this plane the force $A\mathbf{x}$ lies in the plane (so that we can forget about the rest of $\mathbb{R}^n$). And since the work done by $A\mathbf{x}$ around a circle in this plane—centered around the origin—is zero, the tangential component of $A\mathbf{x}$ changes sign at some point(s) $\mathbf{x}_0$ on the circle, which is to say that $A\mathbf{x}_0$ is normal to the circle at $\mathbf{x}_0$. Thus, $\mathbf{x}_0$ is a (real) eigenvector.

3. Orthogonality of the eigenvectors: a physical/geometrical proof. Let $\mathbf{u}$, $\mathbf{v}$ be two distinct eigenvectors of a symmetric $n \times n$ matrix A with the eigenvalues $\lambda \neq \mu = 0$ (the latter assumption involves no loss of generality since we can take $\mu = 0$ by replacing A with $A - \mu I$). Figure 3 shows the force field $A\mathbf{x}$ of such a matrix. Consider the work of $A\mathbf{x}$ around the triangle OQP. The only contribution comes from PO since $A\mathbf{x}$ vanishes along OQ and is normal to QP. And if $A\mathbf{x}$ is conservative, then $W_{PO} = 0$ and hence $P = O$, implying $\mathbf{u} \perp \mathbf{v}$. This completes a "physical" proof of orthogonality of the eigenvectors of symmetric matrices.

4. The entry a_{ij}, $i \neq j$ of a square matrix $A = (a_{ij})$ has a dynamical interpretation: it is the angular velocity, in the (ij)–plane, of $\mathbf{e}_i$ moving with the vector field $A\mathbf{x}$.[1] Indeed, $a_{ij} = (A\mathbf{e}_i, \mathbf{e}_j)$, the projection of the velocity $A\mathbf{e}_i$ onto $\mathbf{e}_j$, Figure 4. And thus the symmetry condition $a_{ij} = a_{ji}$, illustrated in Figure 4, also amounts to stating that the 2D curl in every ij–plane vanishes. The symmetry for 3×3 matrices is equivalent to curl $A\mathbf{x} = 0$. In fact, decomposition of a general square matrix into its symmetric and antisymmetric parts amounts to decomposing the vector field $A\mathbf{x}$ into the sum of a curl-free and divergence-free field, a special case of Helmholtz's theorem, itself a special case of the Hodge decomposition theorem.

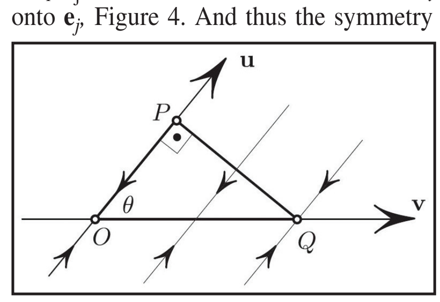


Figure 3. *Geometrical proof of orthogonality.*

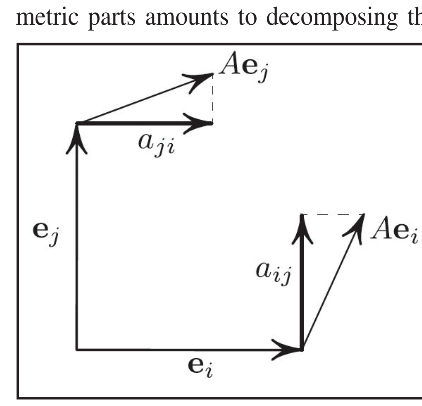


Figure 4. *a_{ij} as the angular velocity.*

And the diagonal entries a_{ii} give the rate of elongation of $\mathbf{e}_i$; this explains geometrically why the cube formed at $t = 0$ by $\mathbf{e}_i$ and carried by the velocity field $A\mathbf{x}$ changes its volume at the rate tr A (at $t = 0$). This also offers a geometrical explanation of the matrix identity $\det e^A = e^{\operatorname{tr} A}$.

[1] to be more precise, we should be referring to the moving vector instantaneously aligned with $\mathbf{e}_i$.

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

Math within March Madness

CAREERS IN MATHEMATICAL SCIENCES

By Tim Chartier

Every year, a fever spreads through much of the United States. It peaks in mid-March and continues through early April. It's called March Madness, the National Collegiate Athletic Association (NCAA)'s Division 1 basketball tournament. The event garners much attention before any time has ticked off a game clock, prior to any ball being shot or dribbled. Millions of people create brackets to predict the outcome of every game in the tournament and to compete against their friends and colleagues for pride, office pool winnings, and sometimes thousands of dollars.

The process of creating a bracket is limited to a few days since matchups are announced less than a week before teams play the first round of 32 games. A completed bracket contains your predicted winner for each of those 32 first-round games. Then you select winners for your 16 predicted second-round matchups. You continue this process until you reach your predicted teams vying for the national championship.

In 2014, Warren Buffett insured a billion dollar prize for anyone who could complete a perfect bracket for the tournament. Correctly predicting the outcome of 63 bracket matchups is not an easy task. In 2014, every one of the over 11 million brackets submitted to ESPN's online tournament had missed a prediction after the first round.

Davidson College basketball team planning their game strategy. Photo credit: Tim Cowie (DavidsonPhotos.com).

For me, this is more of a time of March *Mathness* than Madness. In 2009, along with my collaborator Amy Langville of the College of Charleston and our student researchers, I applied our new ranking research to March Madness. Our research adapted the Colley and Massey methods, used to help rank college football teams and determine which teams play in New Year's Day bowl games. Both methods form a matrix system from the results of a season.

When forming the linear system for the Massey method, let's define the point differential for a single game between a winning team W and a losing team L as $d_{WL} = points_W - points_L$. The Massey method derives its linear system from the assumption that $r_W - r_L = d_{WL}$, where r_W and r_L are the Massey ratings for the winning and losing teams, respectively. Applying least-squares to this over-determined system results in the normal equations $M\mathbf{r} = \mathbf{p}$, where $\mathbf{r}$ is the ratings vector. The Massey matrix can also be formed directly from game data. The diagonal of M results from placing a team's total number of games, t_i, on the diagonal $m_{ii} = t_i$. If g_{ij} represents the number of times teams i and j have played each other, the off-diagonal entries in the Massey matrix are $m_{ij} = -g_{ij}$.

For the Massey method, $\mathbf{p}$ contains information regarding a team's point differential for each game: $p_i = \Sigma_j d_{ij} - \Sigma_k d_{ki}$. The sums over j and k represent the point differential from games that team i has won and lost, respectively. Since the sum over k is subtracted from the sum over j, the sign of each entry in $\mathbf{p}$ gives an indication of a team's scoring performance over the course of a season. Unfortunately, the linear system in $M\mathbf{r} = \mathbf{p}$ has infinitely many solutions for any season. To reduce the number of possible solutions, we add the restriction that all of the ratings sum to zero by replacing the last row of $[M \mid \mathbf{p}]$ with $[1 \ 1 \ \ldots \ 1 \mid 0]$. While infinitely many solutions are still possible, many systems will now have a unique solution.

Another option is to form the linear system for the Colley method, $C\mathbf{r} = \mathbf{b}$, where C is the Colley matrix and $\mathbf{r}$ contains the ratings as a vector. First, $C = M + 2I$. The vector $\mathbf{b}$ contains information regarding each team's number of wins and losses: $b_i = 1 + \frac{w_i - l_i}{2}$. The Colley matrix C is strictly diagonally dominant, so the Colley method produces a unique rating vector r.

Our math-based brackets solve these linear systems and assume a higher-ranked team wins any matchup. In 2010, I was teaching a portion of the methods to undergraduate students at Davidson College, one of whom created a bracket that beat over 99.9% of more than 5 million brackets on ESPN that year.

There are two important pieces to these methods. First, linear systems allow for interdependence of teams' ratings. For example, if you lose to a weaker team, that game hurts you in the standings more than if you lose to a stronger team. This approach integrates strength of schedule into the rating method. Second, the new research allows predictive elements to be weighted. We weighted recency (among other factors), so teams that were playing well going into the tournament were rewarded. The linear systems themselves are created for approximately 350 teams in over 5,000 games, which enables us to find teams that might otherwise be overlooked.

The success of this work got the attention of the media. In 2014 alone, I spoke with *The New York Times*, *USA Today*, and *CBS Evening News* about this topic. There is a madness to March for me now, with much of the month being spent helping the public and my students create brackets.

My work in March Madness launched my research into the field of sports analytics, a subfield of data analytics. Data analytics ideally affects decision-making by studying data, and sports analytics applies this to athletics. By translating scouting notes into matrices that can be analyzed and studied, one can better inform a team about whom to draft, sign, or acquire in a trade. An athlete's performance is studied via hits, rebounds, points, times, or an aggregation of such statistics; often the goal is to find a trend that can increase the odds of winning.

Using my research, I work with over a dozen students to help the men's basketball team at Davidson College. Throughout the season, we sift through a wealth of statistics on basketball databases to scout our opponents. We also record game-to-game statistics not offered on the web for our analysis and study the effect of every lineup coaches use in a game.

My research group also aids professional sports organizations. We work with the NBA using SportVU data. To create these datasets, cameras located in the rafters of every NBA arena record (x, y) coordinates of every player on the court and (x, y, z) coordinates for the ball every 25th of a second throughout the game. While the specifics of our research falls under a nondisclosure agreement, I can share that our studies focus on officiating in one specific instance in a game. Our research supports the NBA league office as it studies the game to ensure officiating is consistent and fair. We have also worked with NASCAR teams, which is natural given that many teams are located near Davidson. Among our various projects, we created an algorithm that would detect loose wheels on a racecar. Our method read time-varying pressures recorded from instrumentation connected to the pneumatic torque gun (impact wrench), which is used during a pit stop to install and tighten nuts on all five bolts in less than 1.5 seconds.

If you're interested in sports analytics, you'll likely find the trajectory of my career compelling. When I entered the job market I had a great postdoc, had done exciting research at national labs, and had many pieces to my professional puzzle in place. Still, I likely would not have been someone's first choice for sports analytics research. At that time, my work was in numerical partial differential equations. Do I model the NBA as a PDE? No. I use linear algebra, the mathematical tool for my work in PDEs. My shift into sports analytics came via ranking methods, and what followed was far from my original plan.

Nevertheless, the shift was very intentional. Upon receiving an Alfred P. Sloan research fellowship, I used the grant in part to move my work from PDEs to ranking. The shift took significant energy but fit my professional goals. For example, ranking, which merged into the larger discipline of sports analytics, was an area in which I could involve more students at earlier stages of their undergraduate studies; over 20 students have worked with me during the last year. We offer analytics to our college teams, professional sports organizations, and businesses with national markets.

I thoroughly enjoy my research field. Still, I can't tell you the stats of any current baseball player. I forget the officials' signals and only know someone fouled or scored. I'll miss a game to play with my kids or walk with my wife, although she isn't always willing to miss a game. What led me into this field is my mathematical ability and leveraging what I do well. I enjoy what I do because I get to see students become independent researchers, I like helping coaches and sports organizations gain insight from data, and I investigate research questions using mathematical techniques that deeply interest me.

I anticipate that I will work in sports analytics for a while, possibly a long while. Yet I don't know for certain, and can enjoy that unknown. Life, in a way, is its own research question. Dive in. Study. Get confused. Discover and keep exploring. I think this allows for a career that isn't a game in itself but can still be defined as winning, even with the inherent madness of life.

Tim Chartier is an associate professor of mathematics and computer science at Davidson College. He received a national teaching award from the Mathematical Association of America and has worked with Google and Pixar on their educational initiatives.

Sue Minkoff (sminkoff@utdallas.edu) of the University of Texas at Dallas is the editor of the Careers in Mathematical Sciences column.

Cambridge Hosts 20th Anniversary Meeting of UKIE SIAM Section

By Nicholas J. Higham

The SIAM United Kingdom and Republic of Ireland (UKIE) Section is celebrating its 20th year. Every January the section holds a one-day annual meeting. Since the 1997 inaugural meeting in Manchester, it has gone from strength to strength and membership has more than tripled to 600 (all SIAM members residing in the UK or Ireland are automatically members of the section).

The 20th meeting was hosted by the Department of Applied Mathematics and Theoretical Physics (DAMTP) at the University of Cambridge on January 7. The event was organised by section officers Des Higham (University of Strathclyde, President), Jennifer Scott (Rutherford Appleton Laboratory, Vice President), Angela Mihai (Cardiff University, Secretary/Treasurer) and local organiser Arieh Iserles (Cambridge). Around 80 people attended, about half of whom were students. As befits an anniversary, the event featured a stellar programme with invited talks from both senior and early career researchers, as well as student posters.

2014 Fields Medalist Martin Hairer (University of Warwick) gave the opening talk. Hairer described how to treat ill-posed stochastic partial differential equations (PDEs) arising in phase transition problems, emphasizing connections between stochastic analysis and numerical methods. Computer animations beautifully illustrated his talk, which was a superb example of how to communicate deep ideas with a minimum of equations.

Simon Tavaré, who is Director of the Cancer Research UK Cambridge Institute and holds appointments in the Department of Oncology and DAMTP at Cambridge, gave a talk entitled "How Useful is Mathematical Modelling in Cancer Research?" Tavaré, who began a two-year term as president of the London Mathematical Society in November 2015, focused on mutations in DNA and what they reveal about cancer evolution. Tavaré explained the role of quantitative methods in understanding how tumours evolve, giving particular emphasis to approximate Bayesian computation. He also offered mathematical scientists advice on where to publish, arguing that collaborative research has the best chance of being noticed by cancer specialists when published in cancer journals.

2014 Fields Medalist Martin Hairer of the University of Warwick gave the opening talk entitled "Stochastic PDEs and Their Approximations." Photo credit: Nicholas Higham.

The UKIE section has been proud to host SIAM presidents at several previous meetings, and did so once again. In her talk, "Models of Transiently Networked Fluids: Wormlike Micelles," SIAM President Pam Cook (University of Delaware) discussed complex fluids that exhibit transiently networked structures at the mesoscale. She described experiments and mathematical models at the macroscale (coupled nonlinear reaction-diffusion equations) and the mesoscale (stochastic differential equations), with motivational examples from the cosmetics, food, and manufacturing industries.

Simon Tavaré of the Cancer Research UK Cambridge Institute addressed the question "How Useful is Mathematical Modelling in Cancer Research?" Photo credit: Nicholas Higham.

In a lecture sponsored by the UK's Institute of Mathematics and its Applications, Barbara Wohlmuth (Technische Universität München) spoke about reducing computational complexity in PDE simulations. Wohlmuth described dimension reduction techniques with the modelling of noise transmission through floors and walls in buildings as one application; the importance of this will be clear to anyone who lives in an apartment with noisy neighbours.

The meeting also featured shorter talks by early career speakers: Jennifer Pestana (Strathclyde) on "Symmetrizing Toeplitz Matrices and Consequences for Solving Linear Systems," Sarah Mitchell (University of Limerick) on "Numerical Challenges Facing an Application of Stefan Problems: Continuous Casting of Metals," and Jasmina Lazic (MathWorks, UK) on "MATLAB and the Mathematics of Our Lives: from Stacking Shelves in Supermarkets to Personalising Car Insurance Premiums."

In the business meeting, Mihai reported on the section's activities over the last year, which included sponsoring speakers and student attendees at relevant meetings in the UK and Ireland. The UKIE section works with the 11 SIAM student chapters in its area, all of which had been invited to send a representative to the section meeting.

Attendees were also invited to a dinner at Magdalene College. The intimate, candlelit setting of The Parlour (at one time the Master's drawing room) was perfect for post-session discussions.

The 2016 meeting realized the hope expressed by the *SIAM News* report of the first-ever UKIE meeting held in 1997: "May it be the first of many SIAM conferences to take place in the UK or the Republic of Ireland." We look forward to many more years of SIAM activities in the UK.

Nicholas J. Higham is Richardson Professor of Applied Mathematics at the University of Manchester and president-elect of SIAM.

Pent Up: Using Pentagons to Tile a Plane

By Casey Mann, Jennifer McLoud-Mann, and David Von Derau

Tilings, or tessellations, refers to a branch of discrete geometry that involves covering the plane (or space) with shapes, and without overlaps. The discipline finds application in a variety of areas, including crystallography, self-assembly, art and design, materials science, biology, and computer graphics, to name a few. Recreational mathematicians enjoy the topic as well, due perhaps to tilings' connection to art and games. Despite the potential application, the field abounds with fundamental open questions reflecting the complexity of the real world that tilings model.

Consider the challenge of understanding which convex polygons give rise to monohedral tilings of the plane (monohedral tilings are those in which all tiles of the tiling are congruent to one another). Quick verification indicates that any triangle or quadrilateral, convex or not, admits tilings of the plane. Skipping the question of pentagons momentarily, the problem of classifying the convex hexagons that tile the plane is solved [5], though the solution is nontrivial. Convex polygons with seven or more sides do not admit any tilings of the plane [1, 4]; this result is also nontrivial and relates to Euler's famous formula for planar graphs. Thus remains the problem of convex pentagons, which may be stated as the following: *In terms of the measures of the angles and sides, classify all convex pentagons that admit monohedral tilings of the plane.*

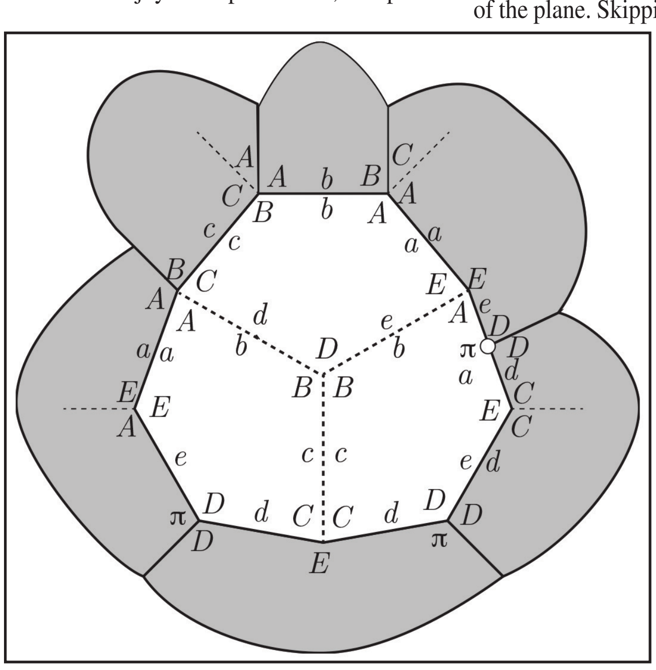


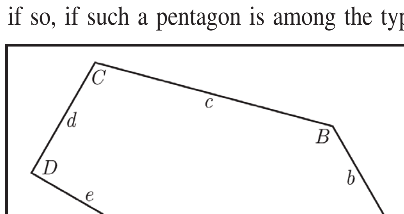
Figure 1. *A hypothetical 3-block with a choice of labeling and a demonstration of how a centrally placed 3-block would be surrounded in a choice of how it might admit a tile-transitive tiling.*

It seems surprising that this simply-stated problem has not been solved, despite a rich history of effort. Attempts date back to Hilbert's famous 23 problems, and include the spotlight of Martin Gardner's *Scientific American* column and notable contributions from amateur mathematician Marjorie Rice [6]. Our team recently made progress on the problem, which we will outline here.

Pentagons admit the most complex monohedral tilings among convex polygons; there are existing types that admit no tilings in which every tile is in the same transitivity class, with respect to the symmetry group of the tiling. This was first demonstrated in 1968 [2]. Theoretically, our work focused on showing that if a pentagon admits tilings with i transitivity classes, there is a maximum number of ways that the tiles of such a tiling can meet one another. This led to the development of a computer algorithm that can, for each positive integer i, exhaustively list all such pentagons [3].

For example, if a hypothetical pentagon admits a tiling of the plane having three transitivity classes, then inside the tiling the pentagon must form into clusters of three pentagons (see Figure 1) so that the cluster of three pentagons tiles the plane in a tile-transitive manner. For such a hypothetical pentagon, we can program a computer to list all labelings of the pentagons comprising this cluster, as well as all the ways such a cluster of three pentagons can tile the plane in a tile-transitive manner. For example, labeling the pentagons in the cluster of three in Figure 1 and requiring it to tile the plane in a specific tile-transitive manner yields a patch of tiles surrounding a centrally-placed cluster of three pentagons, from which we can comprehend relationships among the angles and sides.

From the resulting system of equations, one must then determine if an actual convex pentagon can satisfy this set of equations, and if so, if such a pentagon is among the types already known. In our example, the answer is yes, the equations can form a pentagon, and no, the pentagon is not among the previously-known types. This particular pentagon was the first new type found since 1985 (see Figure 2a above and 2b on page 12).



$A = 60°$	$a = 2$
$B = 135°$	$b = 1$
$C = 105°$	$c = \frac{\sqrt{2}}{\sqrt{3}-1}$
$D = 90°$	$d = 1$
$E = 150°$	$e = 1$

Figure 2a. *Pentagon type 15.*

Moving forward, there is still work to be done on the problem. A classification still eludes us, and more new pentagon types

See **Pentagons** *on page 12*

Comparing Notes on Computing Across the Curriculum

Efforts to Enhance Course Content for Student Exposure

By Kathleen Fowler, Jeffrey Humpherys, Eric Kostelich, Suzanne Weekes, and Lee Zia

At the 2016 Joint Mathematics Meetings, held in Seattle this January, SIAM hosted a panel in collaboration with the American Mathematical Society and the Mathematical Association of America. The panel focused on "Computing Across the Curriculum." Motivated by the emergence of data science, industrial mathematics, and mathematical modeling as necessary workforce skills, faculty are now considering the challenges and benefits of incorporating computing into their courses. Lee Zia, Deputy Division Director for the National Science Foundation's Division of Undergraduate Education (DUE), chaired the session.

To help frame the presentations and subsequent discussion, panelists Kathleen Fowler (Clarkson University), Jeffrey Humpherys (Brigham Young University), Eric Kostelich (Arizona State University), and Suzanne Weekes (Worcester Polytechnic Institute) considered several questions. (1) What does the phrase "computing across the curriculum" mean to you? (2) What forms does "computing" take in this characterization, and in what ways are "data science" or "big data" being addressed? (3) Describe the way(s) in which your department has incorporated computing within your curriculum, and the challenges you have encountered in implementing such a vision. (4) What opportunities to collaborate with other disciplines have you and your colleagues exploited?

Panel members described efforts to reinforce course content by integrating computing into pre-existing courses. Humpherys offered examples of teaching least squares paired with a focus on the numerical methods used to compute solutions, such as QR decomposition. This idea extends to analysis, linear algebra, optimization, probability/statistics, differential equations, control theory, etc., as well as approximation theory and theoretical computer science. Ultimately, students are able to code a simplified algorithm and compare output to an industrial strength version. Once they prove the concept computationally, they can and are expected to use the industrial strength algorithm. Humpherys encourages the use of the term "predictive analytics" over "big data," observing that while the latter term is in vogue, it obscures the important point that mathematical analysis is useful in many settings. Even "small" examples lay the groundwork for understanding general applications of mathematical thinking.

Fowler considers the inclusion of computing components as a necessary method for helping students grow into innovative problem solvers. She requires computing with applied projects in her large-scale freshmen calculus courses. Although grading is a challenge, the trade-offs are worthwhile. Students gain early experience in writing technically, working collaboratively, and tackling open-ended problems. First-year students typically have some experience with Excel, and advanced students often use Python or Matlab, exploring how these programs can be used with modeling to approach real-world problems.

At Worcester Polytechnic Institute (WPI), computing across the curriculum is a priority. All students take calculus, which includes an hour in the computing lab with Maple exercises to reinforce or introduce calculus concepts. Non-faculty instructors and graduate teaching assistants lead these labs. Weekes says that all classrooms are equipped with computer projection systems, so faculty can readily demonstrate concepts using their favorite software. For example, differential equations faculty regularly use Matlab and Maple to demonstrate concepts like resonance and solutions to systems such as predator-prey models and the SIR model.

Weekes also spoke about higher-level courses such as linear programming and math modeling with ODEs, where students solve interesting problems and explore theory after an introduction to Matlab and access to functions such as *linprog.m* or *pplane6.m*. Some faculty use COMSOL and Maple in a lab section of a course entitled Boundary Value Problems to have students solve BVPs numerically and to, for example, demonstrate the collision of solitons. In Probabilistic Methods in OR, a colleague uses Python/NumPy for in-class examples, and students use a programming language of their choice to simulate Markov chains and implement the Metropolis Hastings algorithm.

Weekes emphasized that having excellent computing support resources at WPI has been key to making computing across their curriculum successful. In particular, some university staff offer training sessions to introduce students to scientific software applications; they do this outside of regular class time.

Kostelich suggested that, as motivation, interested faculty look at the January/February 2000 issue of *Computing in Science and Engineering*, edited by Jack Dongarra and Francis Sullivan, which presents a list of the "top 10 algorithms" of the 20th century;[1] the list includes the Metropolis algorithm, the simplex method for linear programming, Quicksort, and the fast Fourier transform. Kostelich would add the Kalman filter, public-key cryptography, and shotgun genome sequencing to their list – but regardless of what one's "top 10" might be, Kostelich argued that all mathematics undergraduates should have some in-depth exposure to a few of them.

The panel was well received and the panelists fielded questions regarding how to advance computing culture in audience members' own home departments. In addition, each panelist spoke of his or her interactions with other disciplines; they all agreed that finding allies in different departments can help increase the number of students taking mathematics, which is ultimately good for home departments. Panel members also urged mathematicians to seek such allies in disciplines like economics and biology, in addition to the more traditional ones in engineering and the physical sciences. Many audience members voiced excitement about the notion of computing across the curriculum, but expressed simultaneous concerns about making it a reality. Conversation indicated that a shortage of graduate students to serve as TAs, lack of support from peers and department chairs, inadequate computing facilities, and outdated course offerings are all real hurdles.

To make a case for integrated computing in mathematics courses, Zia and an audience member also discussed the need for careful assessment. Such assessment could help promote change and demonstrate the benefits and trade-offs in this form of curricular improvement. Zia pointed out that DUE's core funding program, Improving Undergraduate STEM Education (IUSE), is a natural place to seek support for such work. He added that mathematicians have not been as active in submitting proposals as their other disciplinary colleagues, and encouraged the field to engage in such efforts.[2]

To this end, a major takeaway from the session was that efforts are needed to prompt faculty at a wide range of colleges and universities into integrating more computing in their curriculum. The mindset already exists among a majority of SIAM members (and the practices of many SIAM faculty members). With the formation of the SIAG on Applied Mathematics Education (SIAG/Ed),[3] SIAM members have an opportunity to help share best practices, advice, and support with colleagues who have a genuine interest in making these changes. The 2016 SIAM Conference on Applied Mathematics Education, to be held September 30-October 2 in Philadelphia, is an ideal place to generate more discussion. The SIAM Education Committee is in the process of incorporating this theme into some of the proposed minisymposia for the conference.

Kathleen Fowler is an associate professor of mathematics at Clarkson University and a member of the SIAM Education Committee. Eric Kostelich is President's Professor of Mathematics at Arizona State University, where he runs a summer undergraduate research program. Jeffrey Humpherys joined the Department of Mathematics' faculty at Brigham Young University in 2005. He also serves on the SIAM Education Committee, along with Suzanne Weekes, an associate professor in the Department of Mathematical Sciences at Worcester Polytechnic Institute. Lee Zia is the Deputy Division Director for the National Science Foundation's Division of Undergraduate Education (DUE), prior to which he was a professor in the Department of Mathematics at the University of New Hampshire.

[1] An article summarizing the list appeared in Volume 33, Number 4 of *SIAM News*.

[2] http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505082&org=DUE&from=home

[3] http://www.siam.org/activity/ed/

Professional Opportunities

Send copy for classified advertisements to: ***marketing@siam.org****; (215) 382–9800 ext. 364. The rate is $3.15 per word (minimum $395.00). Display advertising rates are available on request. 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****.*

Announcements

Send copy for announcements to: marketing@siam.org; (215) 382–9800 ext. 364. The rate is $2.15 per word (minimum $300.00). For details visit www.siam.org/advertising.

ADVANCE Grants: Narrowing the Gender Gap in STEM Fields

By Heather Walling Doty

SIAM President Pam Cook has worked toward enhancing opportunities for women in academia.

Babatunde Ogunnaike, dean of the College of Engineering at the University of Delaware (UD), says that ignoring the power and strength of diversity is like going through life with one eye closed. Your depth perception is impaired and you don't observe the full richness of what the world has to offer.

SIAM President Pam Cook agrees. Unidel Professor of Mathematical Sciences and associate dean of engineering at UD, Cook fosters the power and strength of diversity in each of her roles. When interacting with students, faculty, and administrators, she aims to create an environment that provides opportunities for those around her to strive to be their best. Cook has had a profound impact on women and underrepresented faculty and students. Nii Attoh-Okine, SIAM member and professor of civil and environmental engineering at UD, says, "Pam's honesty and boldness in tackling the issues of diversity in the academy have made her both a mentor to numerous faculty and the 'go-to person' in critical issues concerning diversity issues in the faculty ranks."

The National Science Foundation (NSF) also believes in the power and strength of diversity. In the year 2000, the U.S. engineering workforce was over 90% male and three-quarters white.[1] Understanding that such homogeneity impaired the nation's potential for scientific and technological innovation, the NSF took action. In 2001, with a goal of diversifying the U.S. science and engineering workforce, it established the ADVANCE program. ADVANCE operates at the faculty level – its initiatives increase the representation and advancement of women faculty in science and engineering. This tactic makes sense; a diverse professoriate will meet the teaching and mentoring needs of a diverse student body, the workforce of the future.

The NSF's ADVANCE program was inspired by actions of the Massachusetts Institute of Technology (MIT) in the 1990s that led to a measurable increase in the number of women faculty in their School of Science. MIT's interventions focused on active recruitment and retention of qualified women faculty. The school strengthened faculty mentoring and ensured equitable allotment of resources. The MIT study, as it has become known, demonstrated that with strong leadership and intentional focus, change is possible.

1 AAUW, *Solving the Equation: The Variables for Women's Success in Engineering and Computing,* American Association of University Women, 2015.

Today, NSF's ADVANCE funds three types of grants. The largest, the Institutional Transformation (IT) grant, is geared toward doing just that – transforming an institution's policies, procedures, practices, and climate to provide opportunity for all faculty to flourish. Social science research on faculty diversity is a required component of IT grants. The other two grants, the IT-Catalyst and the Partnerships or Learning and Adaptation Networks (PLAN), are smaller awards with more targeted scope and function.[2]

Cook is principal investigator on UD's NSF ADVANCE IT grant, a five-year, $3.3 million grant awarded in 2014. UD's IT grant aims to propel women faculty into academic leadership. It puts special emphasis on women faculty post-tenure and women faculty of color.

The grant operates at multiple levels, from the upper administration to the faculty. On the administrative side, the UD ADVANCE team works with the provost's office, providing data-driven talking points – digestible facts and figures on aspects of diversity at UD and/or nationally. UD ADVANCE provides workshops and networking for chairs to help them understand their role in establishing departmental climates and best practices for fair evaluation of faculty, and offers clarification of family-friendly policies and procedures. For faculty, the ADVANCE team offers annual career-development workshops, for example, on the promotion and tenure process for assistant and associate professors and on the path to leadership for full professors. UD ADVANCE is developing mentoring programs specifically for women associate professors in STEM and for women faculty of color.

A decade before UD received its IT grant, the College of Engineering was already doing its part to increase women's representation on the faculty. During this time, the then-dean of engineering hired Cook as associate dean to jumpstart the efforts. Together the dean and faculty made concerted efforts to positively recruit and retain faculty. Supported by a smaller NSF ADVANCE Partnerships for Adaptation, Implementation, and Dissemination (PAID) grant from 2008-2012, Cook led teams of UD faculty in developing workshops on best practices for faculty, conducted by faculty. The workshops were developed in collaboration with faculty from the University of Michigan-Ann Arbor and the University of Wisconsin-Madison, who had established similar workshops through ADVANCE IT grants. The resulting two workshops at UD—one on best practices for faculty recruitment and one on best practices for mentoring faculty—were interactive and included modules on unconscious bias, or the cognitive shortcuts that we all fall back on when we interact with and evaluate others. Workshops were offered annually to faculty search committee members and to senior faculty designated as formal mentors to assistant professors.

2 See the NSF ADVANCE website for details on grant opportunities: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5383.

Pam Cook (third from right), principle investigator (PI) on University of Delaware's ADVANCE Institutional Transformation (IT) grant, along with the ADVANCE team (left to right): Jawanza Keita, Lynn McDowell, Shawna Vican, Joan Buttram, co-PIs Heather Doty and John Sawyer, UD ADVANCE director Emily Bonistall Postel, and co-PI Robin Andreasen. Photo credit: University of Delaware.

The years of focused, collaborative effort paid off at UD, just as they did at MIT in the 1990s. UD's College of Engineering tenured/tenure-track (t/tt) faculty grew from 5% women in 2001 to 17% women in 2011. Of this accomplishment Cook says, "What we learned from the Universities of Wisconsin and Michigan is that workshops by faculty for faculty lead to understanding, buy in, and continued conversations by the faculty after the workshops. And, as the MIT experience showed, diversification takes constant attention and pressure from the administration."

When Cook became a faculty member in the Mathematics Department at UCLA, she was the only female t/tt professor. Another female faculty member was later hired, so that when she received tenure at UCLA she was one of two women t/tt faculty. Upon moving to Delaware (a move precipitated by a spousal employment shift) she was again the only woman t/tt professor. And when she became chair of the Department of Mathematical Sciences, she was the first woman STEM department chair at UD. "A number of us have been solos in our discipline or workspace for much of our lives," Cook says.

Today things have changed at many universities. Family-friendly policies, including parenting leave and stop-the-tenure-clock, are becoming more common. The number of departments with solo women is shrinking, but women faculty in math and science are still underrepresented, especially at the full-professor level. Even at SIAM, an open, inclusive, and international organization, women comprise less than 15% of regular members (among those who have identified their gender). Being such a minority takes its toll. At UD, faculty and administrators continue their work to diversify the faculty, now with the help of the ADVANCE IT award. The goal is to continue to raise awareness and institute policies, and ensure that all faculty and staff are aware of the policies so that the path to advancement is smoother for women, solos, and all faculty.

Pam Cook is current SIAM President and a SIAM Fellow. She is a fellow of the American Association of Science (AAAS) and an associate fellow of the American Institute of Aeronautics and Astronautics (AIAA). Cook's prior positions at SIAM include secretary, vice president for publications, and editor-in-chief of the *SIAM Journal on Applied Mathematics*. She received the national Women in Engineering ProActive Network (WEPAN) University Change Agent award in 2012. Cook's research interests include the mathematical modeling and simulation of fluids. Her early work focused on compressible fluids – transonic aerodynamics, while her current work focuses on viscoelastic (complex) fluids, particularly self-assembling surfactant solutions, mesoscale networked fluids, and gel-like liquids.

Heather Walling Doty is an assistant professor of mechanical engineering at the University of Delaware.

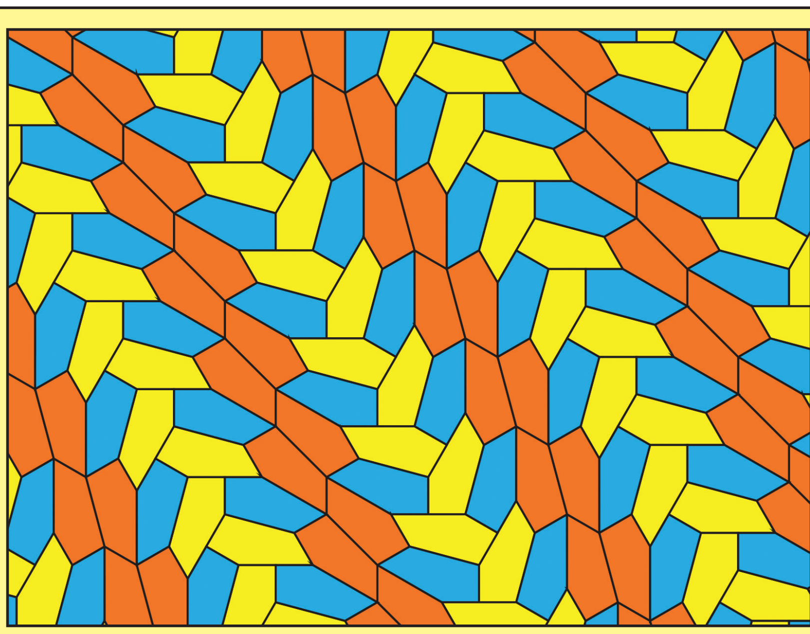


Figure 2b. *Pentagon type 15.*

Pentagons

Continued from page 10

may yet be found. Whether any of these patterns will have direct applications in science is uncertain. However, it seems that furthering our knowledge of how the most basic shapes fit together will be important in understanding the complexity of the real world.

References

[1] Grünbaum, B., & Shephard, G.C. (1987). *Tilings and Patterns*. New York, NY: W.H. Freeman and Company.

[2] Kershner, R.B. (1968). On Paving the Plane. *Amer. Math. Monthly*, *75*, 839–844.

[3] Mann, C., McLoud-Mann, J., & Von Derau, D. (2015). Convex pentagons that admit i-block transitive tilings. *Cornell University Library*. E-print, arXiv:1510.01186.

[4] Niven, I. (1978). Convex polygons that cannot tile the plane. *Amer. Math. Monthly*, *85*, 785–792.

[5] Reinhardt, K. (1918). *Über die Zerlegung der Ebene in Polygone* (Ph.D. thesis) Univ. Frankfurt, Noske (Borna-Leipzig).

[6] Schattschneider, D. (1978). Tiling the plane with congruent pentagons. *Math. Mag.*, *51*, 29–44.

Casey Mann is an associate professor at the University of Washington Bothell. His research interests are in discrete geometry: in particular, tilings and lattice knots. Jennifer McLoud-Mann is an associate professor at the University of Washington Bothell. Her research interests include areas of discrete mathematics such as tiling theory, knot theory, and combinatorics. David Von Derau graduated from the University of Washington Bothell with a degree in mathematics. His research interests are focused on computational modeling and data analytics.