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

siam news

Volume 51/ Issue 3
April 2018

Special Issue on Control and Systems Theory

In this ***special issue****, read about research related to the theory of systems and their control.*

Figure 5. *On the Meuse River, gates move according to feedback laws constructed by Jean-Michel Coron and his colleagues. Photo credit: Jean-Michel Coron.*

In an article on page 6, Paul Davis overviews Jean-Michel Coron's 2017 W.T. and Idalia Reid Prize Lecture, which focused on controllability and local asymptotic stabilizability — two essential properties of control systems.

Robustness of Complex Networks with Applications to Cancer Biology

By Allen Tannenbaum

The study of *complex networks* has a huge and growing literature, and has even been called the field of *network science* [2]. Much of our research is motivated by the need to formulate mathematical principles that are common to various types of complex networks. For example, cell biology, communication webs, and search engines all need to process noisy, uncertain, and incomplete data that is potentially stochastic. We may describe the cell cycle with empirical distributions driven only by partially-known environmental cues and intracellular checkpoints. In turn, search engines employ empirical distributions based on the very limited sampling of features that are often not well understood. This motivates problems that include the characterization of robustness, reliability, and possible uncertainty principles. Our goal is to investigate how curvature and other intrinsic geometric/topological properties affect these features. In short, we want to develop the necessary theory and tools that will permit the understanding and management of network dynamics at various scales.

Network geometry, and curvature in particular, is intimately related to network entropy (see Figure 1, on page 3). In fact, one way to generalize curvature to rather broad metric measure spaces is to exploit entropy's convexity properties along geodesic paths defined at the level of the associated space of probability measures with an induced Riemannian structure.

We are very interested in applications of this geometric network framework, particularly to cancer. The connection to cancer biology arises from one's ability to model many cellular gene and protein networks as weighted graphs, whose edges reflect interaction strengths/rates between the corresponding nodes (genes or proteins). As a concrete example, let us consider a genetic regulatory network. The expression of a gene, i.e., the production of a protein that the given gene encodes, is regulated by other proteins. Thus, one may model the genomic machine as a graph (network), with vertices

See **Cancer Biology** *on page 3*

Sensitive Dependence on Network Structure: Analog of Chaos and Opportunity for Control

By Adilson E. Motter and Takashi Nishikawa

The advancement of network science over the past 20 years has created the expectation that we will soon be able to systematically control the behavior of complex network systems and in turn address numerous outstanding scientific problems, from cell reprogramming and drug target identification to cascade control and self-healing infrastructure development [4]. This expectation is not without reason, given that control technologies have been part of human development for over 2,000 years [1].

While significant progress has been made, our current ability to control is still limited in many systems. This is not so much from lack of available technologies to actuate specific network elements as from challenges imposed by unique characteristics of large real networks to designing *system-level* control actions [4]. These limiting characteristics include the combination of high dimensionality, nonlinearity, and constraints on the interventions, which set networks apart from other systems to which control has been traditionally applied [1]. Recent progress on developing control techniques scalable to large networks has been driven by the design of new approaches.[1]

One such approach may now be possible due to the recent discovery [5] that network dynamics often depends *sensitively* on the network structure, especially when this structure is optimized for maximum stability of a desired behavior. Before considering how this property can be explored to make networks more responsive to control, we discuss in some detail what sensitive dependence on the network structure really means.

When considering a network system's sensitivity to perturbations, one might contemplate whether damage to small parts of the network will compromise the system's structural integrity. This question is meaningful when dealing with a network whose function is primarily structural, such as a spider web, the lattice of crystal materials, or the backbone of a tower. It can be traced back to James Maxwell, whose 1864 work on networks of forces [2] made him a pioneer in the study of networks. Incidentally, this was four years prior to his publication on the flyball governor, a control system then used in steam engines that is often credited as the beginning of the mathematical theory of control [3].

However, there is a potentially much larger class of network systems in which the network's role is not mainly structural. It is instead to mediate a process, as in the case of a road network, power grid, neuronal network, metabolic network, food web, or social network. If one of these networks is perturbed slightly, will the relevant process on the network change only slightly or substantially? In other words, is there a sensitive dependence of the network dynamics on the network structure? For example, in a network of people deliberating an issue, how does the convergence to consensus depend on the details of the network of interactions? In our power-grid network, where collections of power generators must be in pace at approximately 60Hz, how does the

[1] https://sinews.siam.org/Details-Page/leveraging-noise-to-control-complex-networks

See **Network Structure** *on page 3*

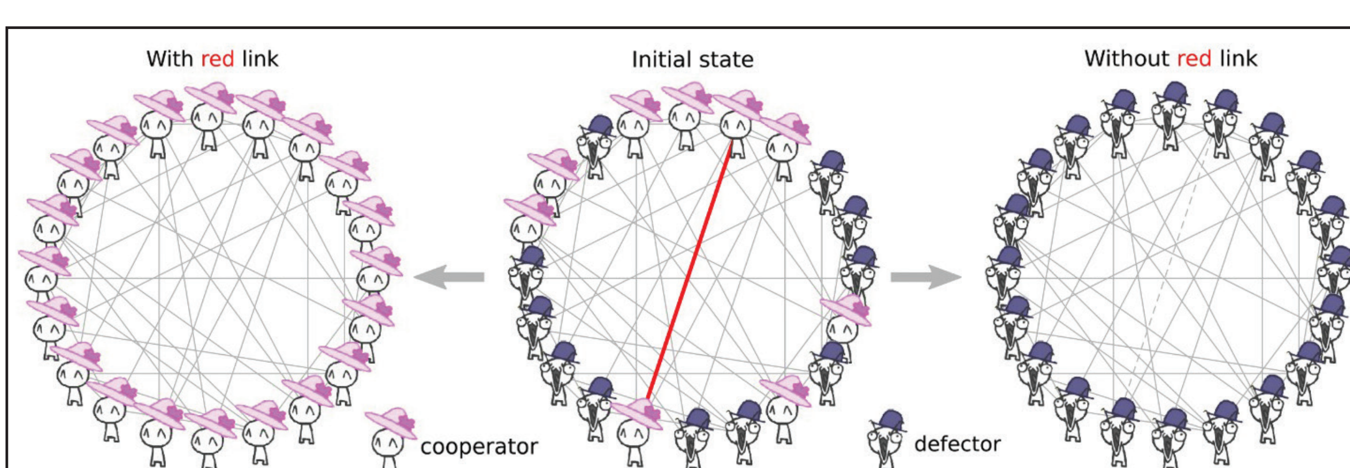


Figure 1. *Final state of two realizations of the iterated prisoner's dilemma game for the same initial condition and the network differing by a single edge. Visuals adapted from Nicky Case's "The Evolution of Trust" game. View an accompanying animation in the online version of this article, or at http://bit.ly/SensPrisDilemma.*

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

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

siam news
Volume 51/ Issue 3/ April 2018

Celebrating *SIAM Review*'s 60th Volume

On page 9, Mark Newman writes about his top-downloaded paper in SIAM Review, which surveyed a rapidly-growing field that continues to flourish today.

Looking Back at 60 Volumes of *SIAM Review*

FROM THE SIAM PRESIDENT

By Nicholas Higham

SIAM was founded in 1952 to promote applied mathematics research and facilitate the exchange of ideas between mathematicians and others in science and industry [3]. Its initial publications were *SIAM Newsletter*, glossy and in U.S. half letter format, and the *Journal of the Society for Industrial and Applied Mathematics* — both of which began in 1953. The newsletter included news about SIAM and the applied mathematics community, book reviews, and adverts. It also contained technical articles of broad interest under the heading "SIAM Notes."

The January 1958 newsletter featured a front-page article announcing a new journal called *SIAM REVUE*,[1] which promised expository papers, book reviews, articles on education and policy, and a problems section.

The first issue of *SIAM Review* appeared in January 1959. Among the articles in early issues were reprints of "SIAM Notes," the first being Thomas N.E. Greville's "The Pseudoinverse of a Rectangular or Singular Matrix and Its Application to the Solution of Systems of Linear Equations." Pseudoinverses proved to be a popular topic in the journal for many years to come.

Beginning with volume 2 of *SIAM Review*, a "News and Notices" section (later renamed "Chronicle") appeared in each issue, intending to "report to the membership the activities of the Society, and news and meeting notices of interest to the applied mathematician." This section took over the role of *SIAM Newsletter*, which was later rejuvenated as *SIAM News* and gradually adopted the task of providing news about SIAM. View archives of *SIAM Review*.[2]

SIAM Review's first issue of 1960 contained a report by Donald Thomsen, Jr. (IBM)—who had just completed his term as SIAM President—on SIAM activities in the preceding year. It described the establishment of a committee to expand SIAM's conference program and the intention to start a SIAM monograph series. The report tallied membership at 1,626, with 47 institutional members and 14 sections. Other officer reports appeared from time to time, including annual reports from the treasurer that from 1986 to 1998 shared much interesting information about SIAM, financial and otherwise. Until 1998, each issue published a list of new members.

In its early years, *SIAM Review* sometimes printed reports of various types of meetings. A 1964 article provides a transcript of a panel discussion on mathematical publishing at the previous year's SIAM fall meeting [4]. It is striking that many of today's issues were also of concern 54 years ago, such as the proliferation of journals, the quality of writing in mathematical papers, and the ways in which editors and referees maintain standards of accepted papers. The comprehensive survey articles for which *SIAM Review* is renowned were not a feature of every issue in the early years, but became more common in the 1970s.

Cartoon created by mathematician John de Pillis.

In the July 1975 issue of *SIAM Review*, the editors announced a new section titled "Classroom Notes in Applied Mathematics" consisting of "brief notes (one to four printed pages), which are essentially self-contained applications of mathematics that can be used in the classroom." They stated that the intention was to eventually collect these notes into a book. This happened in 1987, with the publication of *Mathematical Modelling: Classroom Notes in Applied Mathematics* [7], edited by Murray Klamkin. The same author, who edited the problems section from 1959 to 1993, later published a selection of problems and solutions from the journal [8]. The December 1986 issue featured a paper typeset in TeX that author Layne Watson, an early adopter of TeX, had delivered as camera-ready hard copy [15]. It was not until 1990 that submission instructions mentioned the possibility of authors providing TeX sources for their accepted papers.

The last issue of 1996 announced a major redesign of *SIAM Review*. It arose from recommendations of a 1994 ad hoc committee appointed by the SIAM president to evaluate all SIAM journals. The March 1999 issue was the first to be published in the new format, and a preface from editor-in-chief Margaret Wright explained the changes. A glossy dark blue cover, shiny paper, and a page design unique to the journal were part of the new look. It contained four sections—titled "Survey and Review," "Problems and Techniques," "Education," and "Book Reviews"—with autonomous section editors, along with a "SIGEST" section that reprinted a notable paper of broad interest from another SIAM journal. The previous "Problems and Solutions" section became electronic only. Each section featured an introduction by the relevant section editor that put the articles into context; nowadays these introductions are also posted on SIAM News Online. This format remains essentially unchanged today, though the second section was renamed "Expository Research Papers" in 2010 and "Research Spotlights" in 2012.

Book reviews have always been a popular component of *SIAM Review*, with expert reviewers providing insight into relevant fields and offering due praise and criticism. The redesign introduced "Featured Reviews," in which a reviewer assesses several books on a particular topic. These are especially useful for instructors searching for a suitable classroom text. For example, "PDE Books, Present and Future" by J. David Logan in the September 2000 issue and "Two New Books on Partial Differential Equations" by Ronald B. Guenther and Enrique A. Thomann in the March 2005 issue remain useful guides to textbooks available at those times for a PDE course.

Another recent feature of *SIAM Review* is the publication of reports on educational matters, such as the influential 2001 report on graduate education in computational science and engineering (CSE) [12] (complemented by [16]), which was followed a decade later by a report on undergraduate education [13]. A report on challenges, opportunities, and directions for CSE research and education for the next decade will appear later this year [11].

In preparing this article, I had a lot of fun browsing through the back issues of *SIAM Review*. Here are a few observations. The second most-cited[3] article from the 1960s is actually a half-page problem [14] that introduced a novel application of the orthogonal Procrustes problem. I have even cited it myself. As far as I know, only one paper has had the honor of appearing in *SIAM Review* twice: the famous "nineteen dubious ways" paper on the matrix exponential by Cleve Moler and Charles Van Loan [9]. It was reprinted in 2003 as part of SIAM's 50th birthday celebrations, with new material to bring it up to date [10]. It is interesting to note that some of *SIAM Review*'s most-cited papers have appeared in the Education section, including [2] and [6]. Indeed, six of the 10 most-downloaded papers (see the chart on page 9) appeared in the Education section. Papers with a pedagogic slant can clearly still garner many citations from the research literature. The longest paper I could find (based on some Emacs and MATLAB hacking of the *SIAM Review* BibTeX file in the TeX User Group bibliography archive[4]) is [1], which at 107 pages is longer than some SIAM books. Back issues of *SIAM Review* contain gems that one might not expect to find. One of my favorites is the

See **SIAM Review** *on page 5*

[1] This spelling, which suggests a rather different sort of publication, was only used once in the article and was presumably a typo.

[2] https://epubs.siam.org/loi/siread

[3] All citations are from Google Scholar.

[4] http://ftp.math.utah.edu/pub/tex/bib/

siam news

ISSN 1557–9573.

Cancer Biology

Continued from page 1

representing the genes and edges depicting the correlation (dependence) of a given protein's production by the corresponding gene on additional proteins produced by other genes in the genome.

Accordingly, much of our research focuses on employing network geometrical concepts to quantify (and therefore predict) pathway-related robustness/fragility in a given cancer system. This helps uncover hypothesized sets of targets that can properly disrupt alternative signaling cascades contributing to drug resistance. Our program's mathematical component builds on and relies upon several observations with far-reaching physical (statistical mechanics) and information theoretic significance. More specifically, one can begin by placing a probability structure on a graph (e.g., representing expression levels of genes). This space of probabilities on graphs has several properties that enrich the structure of the underlying discrete space, based on the fact that a Riemannian structure may be endowed on the associated probability measures. Geodesic paths ensue, and convexity properties of the entropy along such paths reflect the space's geometric features.

Entropy's close relation to network topology and robustness has been noted in evolutionary biology literature [3]. For example, Lloyd Demetrius uses Darwinian principles to argue that entropy is a selective criterion that may account for the robustness and heterogeneity of both man-made and biological networks [3]. In our research program, we observe that curvature from network geometry is also strongly related to biological functional robustness. Biological networks seem to display a greater degree of robustness—as exhibited by higher curvature—than random networks.

Based on this observation, we are developing analytical methods for quantitatively describing the functional robustness of cancer networks to identify targets (genes/proteins) of opportunity. We hope that the analytical methods will empower treatments involving targeted drug agents and the combination of immunotherapy with more traditional chemical agents. This will help optimize the efficacy of certain immunotherapy methodologies for the alteration or upregulation of tumor cell antigens. Such an approach involves the use of graph theoretic techniques to identify key cancer hubs by partitioning the network into dense, highly-connected subgraphs. In order to account for both the activation and inhibition properties of the various complex interactions, one must extend existing theory to the case of directed graphs [1].

We are also studying possible mechanisms of resistance. For instance, it seems that the inhibition of certain key pathways (i.e., by making them more fragile) can increase robustness in neighboring pathways and thus contribute to an escape route from a given therapy. Network robustness may also indicate resistance to treatment, while fragility reflects sensitivity. The notion of graph curvature can be quite valuable in quantifying such phenomena. Cancer cells exhibit fate plasticity and are able to shift along a spectrum of differentiation in response to changes in gene expression caused by various genetic assaults (radiotherapy/chemotherapy/immunotherapy) or environmental stresses (hypoxia, reactive oxygen species). The methods we propose can also characterize both the processes that lead to differentiation and targeted anticancer therapies that must account for not only the differentiation state of the tumor as a whole, but also the likelihood that drug-resistant subclones will emerge.

In summary, our work uses ideas from geometric network mathematics in the battle against cancer. Our research is part of the emerging field of *mathematical oncology*, and hopefully will help in the development of new treatments for this deadly disease.

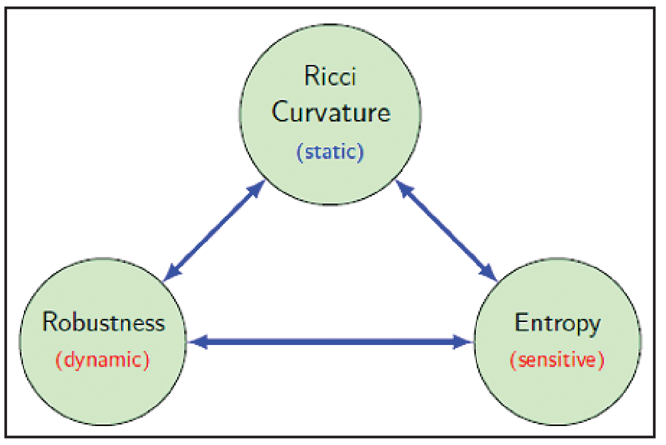


Figure 1. *Network curvature, robustness, and entropy are all positively correlated. Image credit: Liangjia Zhu.*

References

[1] Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. (2006). Boca Raton, FL: Chapman and Hall/CRC Press.

[2] Barabasi, A.-L. (2012). The network takeover. *Nat. Phys., 8*, 14-16.

[3] Demetrius, L.A. (2013). Boltzmann, Darwin and directionality theory. *Phys. Rep., 530*, 1-85.

Allen Tannenbaum is Distinguished Professor of Computer Science and Applied Mathematics & Statistics at Stony Brook University. He has conducted research in algebraic geometry, systems and control, invariant theory, computer vision, medical imaging, and complex networks.

Network Structure

Continued from page 1

stability of this synchronous state depend on the particulars of how the power lines are connected among them? How do the populations of the various species in an ecosystem depend on the specifics of their feeding relationships? And so on.

Our recent study [5], in collaboration with Jie Sun (Clarkson University), examined whether a process can change substantially as a result of small network perturbations *even* when the overall network structure remains uncompromised. For example, consider the evolution of cooperation in an iterated prisoner's dilemma game. Starting with a mixed population of cooperators and defectors, the full population may eventually converge to cooperators only or defectors only, depending on a small structural detail in the network (see Figure 1, on page 1).

To formalize the problem, one can abstract the networks to their graph representations, namely as sets of nodes connected by weighted edges. Of all the networks with a given number of nodes and edges, the most interesting are those evolved or designed to optimize the process under consideration—be it speed to consensus, synchronization stability, or population diversity—where limitations in the availability of edges and nodes generally represent limitations in resource availability.[2]

[2] See the online version of this article, or http://bit.ly/OptNetStruct, for an animation.

Therein lies the rub: as we optimize the network to enhance the relevant dynamical process, this very process may become more sensitive to small changes in network structure. The removal or addition of a node or an edge, or even a small change in edge weights, can cause significant dynamical changes.

As a concrete example, consider network processes governed by the eigenvalues of a coupling matrix. An especially important eigenvalue is the algebraic connectivity—defined as the smallest nonidentically zero eigenvalue of the graph Laplacian matrix—which determines the rate of convergence to a uniform distribution when something diffuses from node to node, the onset of pattern-forming Turing instability in a network, and various aspects of network synchronizability.

We now focus on networks that maximize the algebraic connectivity. As a function of the density of edges, this eigenvalue exhibits cusp-like peaks, which become more pronounced and numerous as the network size increases. In fact, there are infinitely many, infinitely sharp peaks in the limit of large networks (see Figure 2a). Such cusps are signatures of extreme sensitivity to structural changes.

What if we perturb the weights of the edges instead of the number of edges? In some networks, the algebraic connectivity varies linearly with the perturbation strength, while in others it exhibits a very pronounced singular dependence (see Figure 2b). One can determine theoretically a network's sensitivity to edge-weight perturbations, showing in particular that sensitivity is more prevalent in optimal networks.

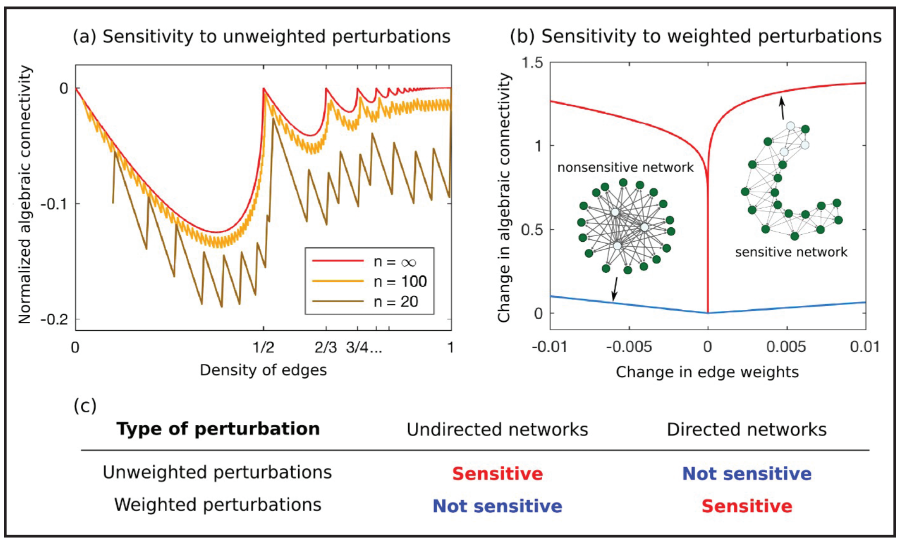


Figure 2. *Sensitivity of optimal networks. The algebraic connectivity can exhibit cusp-like responses to changes in* **(a)** *the density of edges (unweighted perturbations) and* **(b)** *edge weights (weighted perturbations), as summarized in* **(c)**. *Image credit: Takashi Nishikawa and Adilson Motter.*

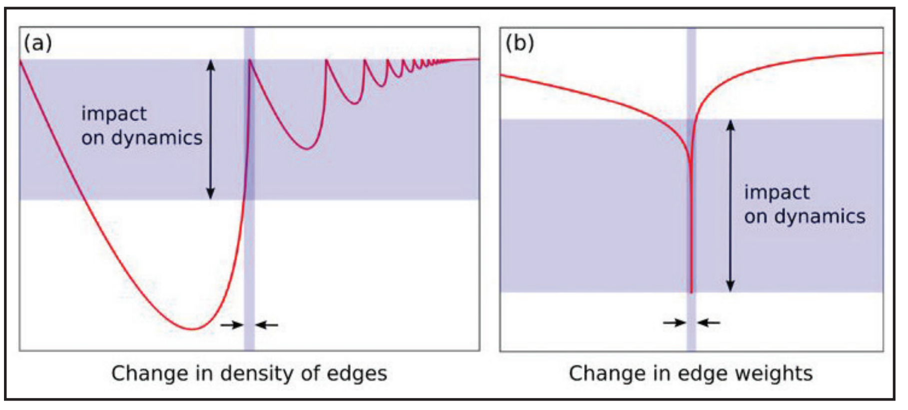


Figure 3. *Control implication of sensitive dependence on network structure. Large changes in the dynamics can be induced by small adjustments in* **(a)** *density of edges or* **(b)** *edge weights. Image credit: Takashi Nishikawa and Adilson Motter. View an accompanying animation in the online version of this article, or at http://bit.ly/NetwControl.*

More broadly, whether the system exhibits sensitive dependence on network structure depends on the class of networks, type of perturbation, and dynamical process. For processes governed by the algebraic connectivity, undirected networks are sensitive to edge removal and node addition but not to edge-weight perturbation, whereas directed networks are sensitive to edge-weight perturbation but not to changes in the number of edges or nodes (see Figure 2c).

The sensitive dependence of collective dynamics on network structure is analogous to the butterfly effect observed in the phenomenon of chaos. The butterfly effect commonly refers to sensitive dependence on initial conditions, where small changes in the initial state lead to large changes in the system's subsequent evolution. Here, large changes in the dynamics are instead determined by small changes in the system's parameters, which in this case define the underlying network. Thus, one can interpret this phenomenon as a parameter counterpart of the sensitive dependence observed in low-dimensional systems. While the effect is associated with chaos in low-dimensional systems, it is induced by optimization in network systems.[3]

How can sensitive dependence on the network structure benefit control? This sensitive dependence to changes is akin to an instability, such as those explored in jet design for increased maneuverability. Moving the center of gravity aft reduces an airplane's stability, and moving it past the neutral point makes the airplane unstable, which increases its response to a given action. In this partial analogy, sensitivity to network structure means that one can manipulate the dynamics substantially with small structural adjustments, as shown in Figure 3. That is, a sensitive network can be responsive to control even when the control actions are highly constrained, either with respect to the number of network components that can be actuated or the extent to which they can be changed. Though limited by how well one can resolve the cusp structure in practice, this property has the potential to lead to new control approaches based on modification of the network's effective structure in real time.

The interplay between network structure, optimization, and sensitivity is a promising topic of future research that offers fundamental insights into the control properties of complex network systems.

[3] See the online version of this article, or http://bit.ly/NetAnalogChaos, for an animation.

References

[1] Astrom, K.J., & Kumar, P.R. (2014). Control: A Perspective. *Automatica, 50*, 3.

[2] Maxwell, J.C. (1864). On Reciprocal Figures and Diagrams of Forces. *Philos. Mag., 27*, 250.

[3] Maxwell, J.C. (1868). On Governors. *Proc. R. Soc. Lond., 16*, 270.

[4] Motter, A.E. (2015). Network controlology. *Chaos, 25*, 097621.

[5] Nishikawa, T., Sun, J., & Motter, A.E. (2017). Sensitive Dependence of Optimal Network Dynamics on Network Structure. *Phys. Rev. X, 7*, 041044.

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

Singular Perturbations in Noisy Dynamical Systems

By Paul Davis

Idealized models are commonplace in scientific theory, though often they are not quite accurate. In some cases, the reality of an almost-but-not-quite-ideal model differs radically from the unattainable ideal model, like in such essential phenomena as chemical reactions. The source of the difference may be the effects of noise, which is frequently modeled as a random perturbation of a deterministic dynamical system.

In his John von Neumann Lecture at the 2017 SIAM Annual Meeting, held last summer in Pittsburgh, Pa., Bernard Matkowsky, John Evans Professor of Engineering Sciences and Applied Mathematics at Northwestern University, surveyed "Singular Perturbations in Noisy Dynamical Systems." As one example, he showed how these tools can quantify the radical difference between the behavior of a particle in a potential well and one that is also subject to collisions with smaller, lighter particles, as in Brownian motion.

In the absence of collisions, the particle is trapped in the potential well. However, collisions with the particles comprising the medium through which the Brownian particle travels will eventually force it from the well, even if the strength of each individual collision is very small. The tools of matched asymptotic expansions (MAEs), and extensions thereof, can predict the expected time to exit and the probabilities of exit locations on the boundary (rim) of the well. The analysis explores the variation of such physical outcomes with the strength of an individual collision.

The dramatic difference between the ideal model and the one with small noise effects included is symptomatic of a singular perturbation. The cumulative effect of random collisions can overcome even the powerful pull of a potential well that is constraining the particle to stay in the well, though each collision results only in an extremely small movement of the particle.

Small terms with big effects call for a singular perturbation analysis. Matkowsky's primary tools are MAEs and extensions thereof. Matching connects the so-called outer solution to the rapidly varying inner, or boundary-layer, solution. The asymptotics involve the limit as a measure of the strength of each collision—typically denoted by a small parameter ε—goes to zero. The two expansions are finite-term approximations in ε.

Matkowsky and his colleagues have modified or augmented MAEs to answer two central questions: (1) "What is the mean time for a particle to escape from a given starting point?" and its spatial counterpart, (2) "What is the mean probability of escape locations on the boundary given the particle's starting point?" Matkowsky's models are typically stochastic differential equations, mathematical representations of deterministic dynamical systems perturbed by small white noise.

The answer to these central questions are asymptotic expansions in terms of the small parameter ε, which depends, e.g., on physical quantities such as temperature, the height of the potential barrier that must be overcome to escape, Boltzmann's constant, etc. Hence, the results connect the observable outcomes of experiments to fundamental material properties.

A short but elegant path built on Ito calculus connects the stochastic dynamical system to two singularly perturbed elliptic boundary-value problems whose solutions capture the expectations of interest. The solution of a Poisson problem with zero boundary conditions gives the mean first passage time to the boundary, while the mean distribution of exit points is the Green's function of a Dirichlet problem with given boundary conditions.

For ease of exposition, Matkowsky considers the problem in one dimension on the interval $(-a, b)$ with $a,\ b>0$, and $x=0$, a stable equilibrium point. The well corresponds to the potential $V=x^2/2$, so that the deterministic force is $-x$. These deterministic problems are singularly perturbed since the diffusion term modeling the collisions is small compared to the potential force. Setting ε to zero reduces the differential equation from second-order to first-order, leaving more boundary conditions to satisfy than degrees of freedom in the outer solution.

One must account for diffusion in the so-called boundary layer. Since the solution varies rapidly there, Matkowsky stretches the spatial coordinate to find the so-called inner solution that satisfies both a second-order equation and the boundary condition. Finally, he matches the two solutions so that they connect smoothly. The final asymptotic approximation is the sum of the outer and inner expansions less their common parts (so that they are not counted twice).

See **Noisy Dynamical Systems** *on page 5*

Untangling Random Polygons and Other Things

By Charles F. Van Loan

The following is an introduction to the 2018 John von Neumann Lecture, to be presented at the upcoming SIAM Annual Meeting (AN18) in Portland, Ore., from July 9-13. Look for feature articles by other AN18 invited speakers introducing the topics of their talks in future issues.

The *polygon averaging problem* involves just about everything I like with respect to teaching and research in matrix computations: my favorite matrix decomposition (the singular value decomposition), my favorite matrix dimension $(N=2)$, and my favorite structured matrix (the "upshift matrix"). Even better, it evolved from an assignment that I gave in a MATLAB-based CS1 introduction, a reminder of just how much there is to learn when working at that level.

Here's the problem: suppose we have a random polygon $\mathcal{P}_0$ with vertices $(x_1,y_1),\ldots,(x_n,y_n)$. Assume that $\mathcal{P}_0$ has centroid $(0,0)$, and $\|x\|_2=1$ and $\|y\|_2=1$. If we connect the midpoints of its edges we obtain a new polygon, also with centroid $(0,0)$. The "next" polygon $\mathcal{P}_1$ is obtained by scaling the x- and y-values so that $\|x\|_2=1$ and $\|y\|_2=1$. The process can obviously be repeated to produce a sequence of polygons $\{\mathcal{P}_k\}$. The assignment required students to plot a reasonable number of $\mathcal{P}_k$. They observed something truly amazing. No matter how "criss-crossy" the initial polygon $\mathcal{P}_0$, the $\mathcal{P}_k$ eventually "untangle" and their vertices head towards an ellipse with a 45-degree tilt (see Figure 1).

This astonished me as well, because I did not work out the solution before handing out the assignment! But that is precisely why I like teaching CS1 courses — they are both an integral part of STEM education and full of surprises. Teaching at the CS1 level allows me to build intuition for linear algebra. The one-dimensional array—a vector in MATLAB—typically marks the first time that a beginner STEM student sees n things as one thing. Yes, a polygon is defined by its n vertices, but a student must think at the vector level when writing the function

```
[xTilde,yTilde]=PolygonAve(x,y),
```

which produces a new, midpoint-connected polygon from the old one. Behind the scenes is a sparse matrix-vector product, e.g.,

$$\begin{pmatrix}(x_1+x_2)/2\\(x_2+x_3)/2\\(x_3+x_4)/2\\(x_4+x_1)/2\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1&1&0&0\\0&1&1&0\\0&0&1&1\\1&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}.$$

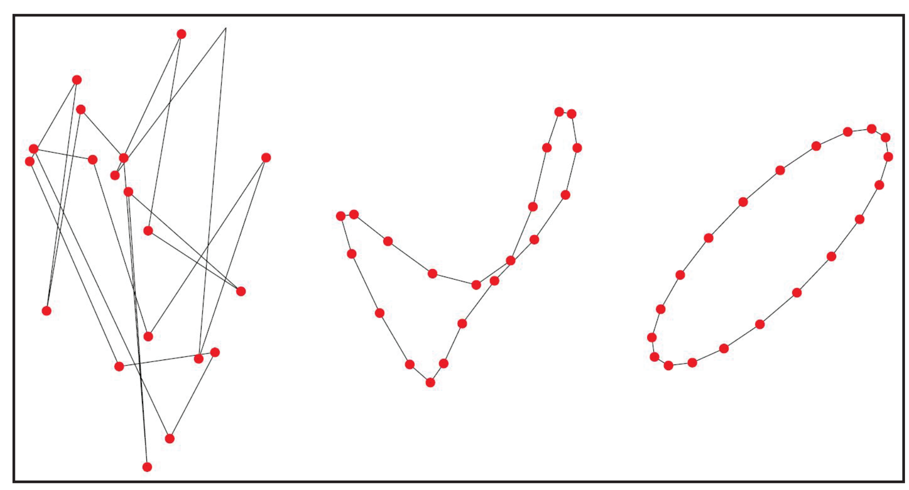


Figure 1. *The progression from $\mathcal{P}_0$ to $\mathcal{P}_{18}$ to $\mathcal{P}_{200}$ for an $n=20$ example. Figure courtesy of [1].*

My CS1 freshmen did not know enough linear algebra to think like this. But they did know enough MATLAB to watch the polygons untangle and ask the ultimate research-driven question: why?

On to the next course, a CS2 treatment of introductory matrix computations. Those students did know enough linear algebra to understand that the transition from one polygon to the next involves a pair of matrix-vector updates. In particular, the x and y vertex updates involve matrix-vector products $x\leftarrow Mx$ and $y\leftarrow My$, where $M=(I+S)/2$ and S is the upshift matrix, e.g.,

$$S=\begin{pmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{pmatrix}.$$

With this notation, one can describe the polygon sequence's production as two separate instances of the power method. One produces the vector sequence $\{M^k x_{orig}\}$ and the other produces $\{M^k y_{orig}\}$, where x_{orig} and y_{orig} house the vertices of the original polygon $\mathcal{P}_0$. Because we know closed formulas for the eigenvalues and eigenvectors of the downshift matrix S (and hence M), we can develop closed-form expressions that completely specify the vertices for each and every polygon $\mathcal{P}_k$. Hint: they involve lots of sines and cosines. The CS2 polygon averaging assignment revolved around building insight for power method convergence associated with the matrix M.

The task then transitioned to a research project with Adam Elmachtoub, an undergraduate at Cornell University at the time. We determined how and why the polygon vertices moved towards an ellipse with 45-degree tilt. It turns out that the invariant subspace associated with the second- and third-largest eigenvalue of $M=(I+S)/2$ is critical to the analysis. The semiaxes and 45-degree tilt of the target ellipse follow from a singular value decomposition (SVD) analysis of a 2-by-2 matrix whose entries are simple functions of x_{orig} and y_{orig} (the singular values and left singular vectors of this 2-by-2 totally specify the tilted ellipse). Full details are available in the resulting *SIAM Review* paper [1].

The path from CS1 to *SIAM Review* was interesting from start to finish. We observed a phenomenon through experimentation and followed it up with a matrix-vector description of that phenomenon and an SVD/eigenvalue analysis that explained everything. Things do not always work out this nicely. Nevertheless, it is fun to think about the polygon averaging problem as simply a metaphor that speaks to the power of matrix-based scientific computing.

I will talk about this trajectory of our work at the John von Neumann Lecture at the 2018 SIAM Annual Meeting.

References

[1] Elmachtoub, A.N., & Van Loan, C.F. (2010). From Random Polygon to Ellipse: An Eigenanalysis. *SIAM Rev., 52*, 151-170.

Charles F. Van Loan is the 2018 recipient of SIAM's flagship John von Neumann Lecture. He is professor emeritus in the Department of Computer Science at Cornell University.

SIAM Review

Continued from page 2

delightful glimpse into the life of Leonhard Euler by Walter Gautschi [5].

Two measures of *SIAM Review*'s influence are its citations and its impact factor. *SIAM Review* is consistently among the top-ranked journals in applied mathematics by impact factor, and often holds the leading position. Many of its papers boast thousands of citations. Table 1 shows the five most-cited papers as of January 2018, according to Google Scholar.

Rank	Author(s)	Title	Year	Number of Citations
1	Mark E. J. Newman	The Structure and Function of Complex Networks	2003	16,721
2	Scott Shaobing Chen, David L. Donoho, and Michael. A. Saunders	Atomic Decomposition by Basis Pursuit	2001	10,396
3	Peter W. Shor	Polynomial-time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer	1999	7,055
4	Benoit B. Mandelbrot and John W. Van Ness	Fractional Brownian Motions, Fractional Noises and Applications	1968	6,952
5	Aaron Clauset, Cosma Rohilla Shalizi, and Mark E. J. Newman	Power-law Distributions in Empirical Data	2009	5,792

Table 1. *The five most-cited papers in SIAM Review per Google Scholar, as of January 2018.*

The mix of older and more recent papers in the table shows that *SIAM Review* not only covers enduring topics but also identifies new and emerging areas. Indeed, the June 2018 issue will present a survey on the very timely topic of optimization methods for machine learning.

SIAM Review's first 60 volumes are a microcosm of applied mathematics and computational science from 1959 until now. As such, they provide a valuable historical record, both in the technical content and the news items about early SIAM activities. However *SIAM Review* evolves over the next several decades, it will remain a must-read for those who want to keep up with innovations in research and education in our field.

References

[1] Asch, M., Kohler, W., Papanicolaou, G., Postel, M., & White, B. (1991). Frequency Content of Randomly Scattered Signals. *SIAM Rev., 33*(4), 519-625.

[2] Berrut, J.-P., & Trefethen, L.N. (2004). Barycentric Lagrange Interpolation. *SIAM Rev., 46*(3), 501-517.

[3] Block, I.E. (1973). SIAM — Its First Three Years. *SIAM Rev., 15*(2), v-ix.

[4] Block, I.E., & Drenick, R.F. (Eds.). (1964). A Symposium on the Publication of Mathematical Literature. *SIAM Rev., 6*(4), 431-454.

[5] Gautschi, W. (2008). Leonhard Euler: His Life, the Man, and His Works. *SIAM Rev., 50*(1), 3-33.

[6] Higham, D.J. (2001). An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Rev., 43*(3), 525-546.

[7] Klamkin, M.S. (Ed.). (1987). *Mathematical Modelling: Classroom Notes in Applied Mathematics*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

[8] Klamkin, M.S. (Ed). (1990). *Problems in Applied Mathematics: Selections from SIAM Review*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

[9] Moler, C.B., & Van Loan, C. (1978). Nineteen Dubious Ways to Compute the Exponential of a Matrix. *SIAM Rev., 20*(4), 801-836.

[10] Moler, C.B., & Van Loan, C. (2003). Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five years Later. *SIAM Rev., 45*(1), 3-49.

[11] Rüde, U., Willcox, K., McInnes, L.C., De Sterck, H., Biros, G., Bungartz, H.,…,Woodward, C.S. (2018). Research and Education in Computational Science and Engineering. To appear in *SIAM Rev., 16*(2).

[12] SIAM Working Group on CSE Education. (2001). Graduate Education in Computational Science and Engineering. *SIAM Rev., 43*(1), 163-177.

[13] SIAM Working Group on CSE Undergraduate Education. (2011). Undergraduate Computational Science and Engineering Education. *SIAM Rev., 53*(3), 561-574.

[14] Wahba, G. (1965). Problem 65-1, A Least Squares Estimate of Satellite Attitude. *SIAM Rev., 7*(3), 409.

[15] Watson, L.T. (1986). Numerical Linear Algebra Aspects of Globally Convergent Homotopy Methods. *SIAM Rev., 28*(4), 529-545.

[16] Yasar, O., & Landau, R.H. (2003). Elements of Computational Science and Engineering Education. *SIAM Rev., 45*(4), 787-805.

Nicholas Higham is Royal Society Research Professor and Richardson Professor of Applied Mathematics at The University of Manchester. He is the current president of SIAM.

Noisy Dynamical Systems

Continued from page 4

The standard rubric of MAE fails for problems exhibiting boundary-layer resonance, such as the one under consideration, because there are not enough conditions to uniquely determine the solution (a constant remains undetermined). This process yields a one-parameter family of possible solutions, though only one can be the actual solution. Matkowsky and his colleagues have developed four approaches to rescue the MAE method and offer deeper insights into the asymptotics.

Two of the approaches impose an extra condition to select a single solution from this one-parameter family. The first chooses the solution that is a stationary point of the Euler-Lagrange variational principle associated with the given boundary-value problem. The second replaces the variational condition with an appropriate orthogonality condition. One may think of the former as related to the Ritz method and the latter as related to the Galerkin method.

A third approach constructs the boundary-layer function in a different way. Rather than an exponent, which is linear in the stretched variable, Matkowsky employs the Jeffreys-Wentzel-Kramers-Brillouin (JWKB) method, which allows the exponent to be nonlinear; in this case, one boundary-layer function can describe two distinct boundary layers. Finally, a fourth approach employs asymptotics beyond all orders, adding an exponentially small term to the outer expansion. The results of the four approaches are the same. Namely, exit occurs through the left (right) end point if $a<b$ $(b<a)$, and exit is equally likely to occur at either end point if $a=b$.

With this panoply of tools, one can then explore the detailed behavior of such fundamental physical processes as chemical reactions and atomic migration in crystals, among others. The important physical observable, e.g., the reaction rate, can be modeled as the random escape of a particle from a potential well. Specifically, the rate is half the inverse of the mean escape time. The factor "one half" enters because the particle is equally likely to exit or return to the well once it reaches the boundary.

Matkowsky's lecture is available from SIAM either as audio or as a PDF of his slides.[1] Alternatively, one can access it on Matkowsky's departmental website.[2]

Since 1960, SIAM has annually recognized a John von Neumann lecturer for outstanding and distinguished contributions to the field of applied mathematical sciences, and effective communication of these ideas to the community. The award honors John von Neumann (1903-1957), one of the most prolific and articulate practitioners of applied mathematics in the 20th century. "The von Neumann Lecture is particularly meaningful to me since four of the previous awardees were my teachers and inspirations," Matkowsky said. "These include Kurt Friedrichs, Peter Lax, Jurgen Moser, and most importantly to me, my advisor and friend Joe Keller."

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

[1] https://www.pathlms.com/siam/courses/4988/sections/7416.

[2] http://people.esam.northwestern.edu/~matkowsky/

Bernard Matkowsky (Northwestern University) delivered the John von Neumann Lecture at the 2017 SIAM Annual Meeting, held last summer in Pittsburgh, Pa. SIAM photo.

Controllability and Local Asymptotic Stabilizability of Control Systems

By Paul Davis

Does SIAM award a prize for balancing a walking cane better than Jimmy Stewart, stabilizing a baby carriage (*sans* baby), and leveling surges in the Sambre and Meuse rivers? Indeed it does, for those are but three of many examples illustrating the profound work for which Jean-Michel Coron received the W.T. and Idalia Reid Prize in Mathematics at the 2017 SIAM Annual Meeting, held in Pittsburgh, Pa., last summer.

Coron is currently a full professor at Sorbonne Université (Paris 6) and a member of the French Academy of Sciences. The Reid Prize citation specifically recognized the importance of "the Coron return method for feedback stabilization of nonlinear systems using time-varying controls," and his prize lecture offered an artful and entirely modest tour of this approach and other substantial contributions.

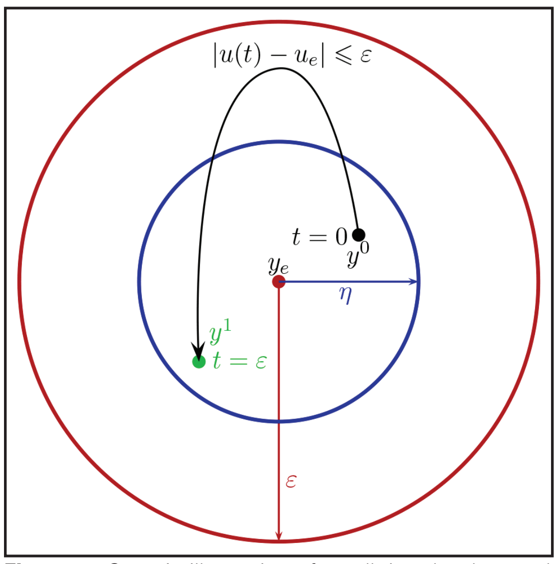


Figure 1. *Coron's illustration of small-time local controllability. The nonlinear control system has an equilibrium at* (y_e, u_e)*; the initial state* y^0 *and the target state* y^1 *are very close to* y_e*; the state remains close to* y_e*; the control remains close to* u_e*; and the time is small. Figure credit: Jean-Michel Coron.*

Coron's work centers on two essential properties of control systems: controllability, or the existence of a control strategy that guides a system from one specified state to another, and local asymptotic stabilizability, or the existence of a feedback law—a control depending on the state—that imparts asymptotic stability to the corresponding closed-loop system. He offered the following example to illustrate the importance of stability. A satellite's orientation can be controlled if it has two or more rocket motors. But the satellite loses local asymptotic stabilizability when only two motors are functioning, and soon drifts out of the desired orientation. Coron noted dryly that the mathematical outcome—the absence of a stabilizing feedback law—appears to warrant national news coverage if the satellite is sufficiently expensive or important.

Coron illustrated the breadth of his work with a variety of applications, both whimsical and practical. He opened his talk with a video displaying feedback control of a moving cart stabilizing an inverted double pendulum into perfect vertical rigidity.[1] This was followed by a saucy clip from Alfred Hitchcock's *Vertigo* in which Jimmy Stewart struggles to balance a vertical walking cane—a mere single pendulum—on his palm.[2]

Although Stewart's balancing efforts did not inspire confidence, the rigidly-vertical inverted double pendulum left no doubt as to its stabilizability. Previewing both the foundation of his own work and what he calls "Louis Nirenberg's advice to depressed mathematicians: Have you tried to linearize?", Coron demonstrated the controllability of the inverted pendulum system linearized around its vertical equilibrium. If a linear system is controllable, a pole-shifting argument reveals the presence of a linear feedback control that renders its zero state globally asymptotically stable. From this point, one can argue that a nonlinear system with a controllable linearization is both small-time locally controllable (Figure 1 illustrates this notion of controllability) and locally asymptotically stabilizable. Stewart's problem with balancing a cane vertically is entirely his own.

Unfortunately, a continuous feedback law cannot locally stabilize all small-time locally controllable systems. Coron's methodical dissection by example cleanly separated the two properties—controllable and stabilizable—and demonstrated the advantage of *time-dependent* feedback for stabilization.

A baby carriage is an example of a controllable system that cannot be stabilized by a continuous feedback law. This system is small-time locally controllable at the origin, but cannot be asymptotically locally stabilized because it does not satisfy Roger Brockett's necessary condition for local asymptotic stabilizability. However, a time-dependent feedback law (see Figure 2) can restore local asymptotic stability of the carriage. Coron cautioned that his analysis is limited to the carriage *without* a baby; with a baby, he could not guarantee small perturbations!

Similar phenomena manifest themselves in both a satellite with less than three functioning thrusters and a quadcopter confined to a plane (also known as a slider). Both eventually crash because they are small-time locally controllable but not locally asymptotically stabilizable by means of continuous feedback laws. Again, time-dependent feedback laws save the day.

Motivated perhaps by his well-placed admiration for Nirenberg, many of Coron's theoretical contributions can be seen as remedies for the gaps and limitations of linearization. His eponymous return method, for instance, avoids a potentially embarrassing failure: what if one could say nothing about the controllability of a nonlinear system near an equilibrium point when the linearized version is *not* controllable? Lie brackets offer an alternative tool in finite dimensions, but can fail for many important partial differential equations (PDEs).

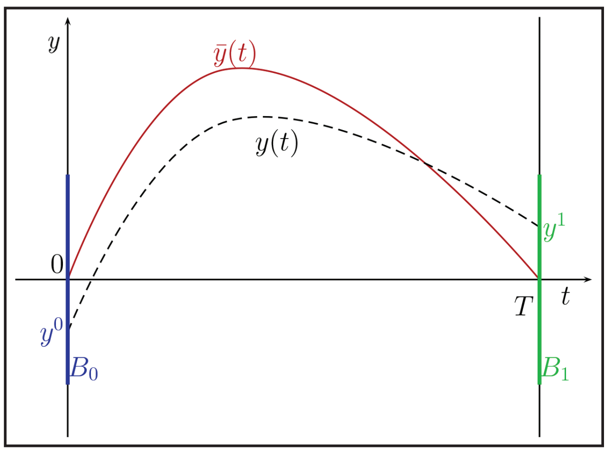


Figure 3. *The idea underlying Coron's return method for assessing controllability of nonlinear systems whose linearization about the equilibrium of interest is not controllable: linearize instead about a nearby trajectory (red) that leaves from and returns to the troublesome equilibrium. Figure credit: Jean-Michel Coron.*

Coron's return method examines the nonlinear system when it is linearized around nontrivial trajectories that begin and end at the problematic equilibrium and for which the linearized system *is* controllable. One can retrieve local controllability of the nonlinear system via an inverse mapping theorem argument (see Figure 3). With such trajectories in hand, the tools of linear control suffice to show that systems like the baby carriage are indeed controllable.

Three quick visuals, two of them quite unassuming, summarized Coron's substantial achievements in understanding the control of PDEs: an animation showing rapid attenuation of one-dimensional shallow water waves in a pool by controlled motion of hydraulic gates at the two ends of the pool[3] (see Figure 4); a photo of gates on the Meuse River moving according to feedback laws constructed by Coron and his colleagues (see Figure 5, on page 1); and a cross-sectional view of the control devices themselves — an adjustable sluice gate and its partner, an adjustable spillway (see Figure 6).

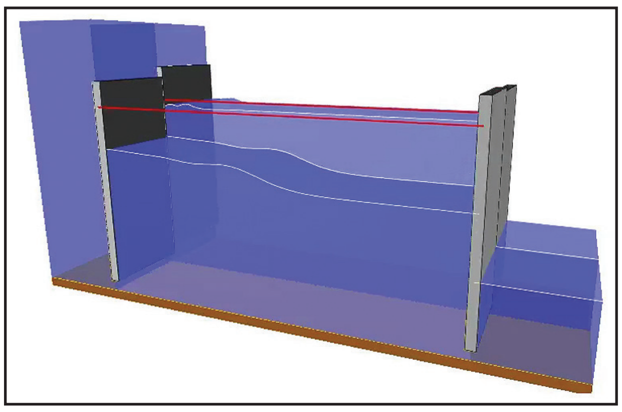


Figure 4. *Water flows below the sluice gates (see Figure 6 for schematic). The back of the figure shows the gates moving up and down according to feedback laws created by Jean-Michel Coron and his colleagues. In the front of the figure, the gates are motionless. A much faster convergence to the desired height of the water, represented by the red line, is achieved by feedback laws. Image credit: Jonathan de Halleux.*

These examples help describe Coron's success in controlling the Saint-Venant shallow water equations. He modestly drew the audience's attention to Adhémar Jean Claude Barré de Saint-Venant's derivation of those PDEs at the age of 74, rather than to his own subsequent accomplishments with them. And Coron barely mentioned his other successes with such fundamentals of fluid mechanics as the Euler equations and the Navier-Stokes equations.

Mirroring the breadth of his analytic contributions, the practical reach of Coron's research goes far beyond movie clips and animations of suppressed waves. He and his collaborators have implemented a stabilizing control strategy for the Sambre and Meuse rivers in Belgium. Coron began working on this subject in 2003, and his feedback laws were implemented a few years later on the Sambre and only recently on the Meuse.

Coron's Reid Prize presentation is available from SIAM as slides with synchronized audio or as a PDF of slides only.[4]

The Reid Prize began with awards by SIAM in 1994, 1996, and 1998, all funded by John Narcisco, nephew of Idalia Reid. Since 2000, SIAM has awarded the prize annually with support from a bequest from Idalia Reid in memory of her late husband, William T. Reid. W.T. Reid worked in differential equations, the calculus of variations, and optimal control, sharing naming rights for the workhorse Gronwall-Reid-Bellman inequality. He held faculty appointments at the University of Chicago, Northwestern University, the University of Iowa, the University of Oklahoma, and the University of Texas. He was an important figure in the optimal control community and a beloved mentor to his students.[5]

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

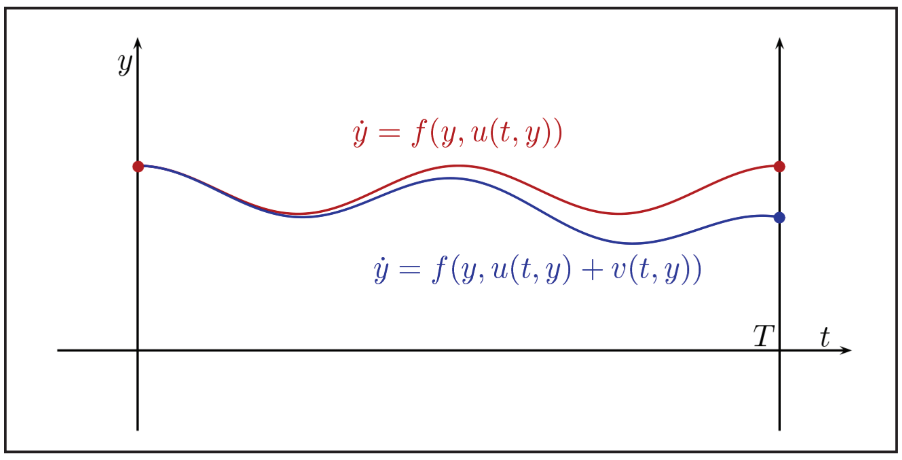


Figure 2. *The idea underlying Coron's return method to stabilize the empty baby carriage—and more generally, driftless control systems—is to construct a preliminary T-periodic feedback* $u(t, y)$*, leading to T-periodic trajectories with controllable linearizations. Using these linearizations, one can then assemble a T-periodic perturbation* $u(t, y) + v(t, y)$ *of* $u(t, y)$*, such that the trajectories are now converging to 0. Figure credit: Jean-Michel Coron.*

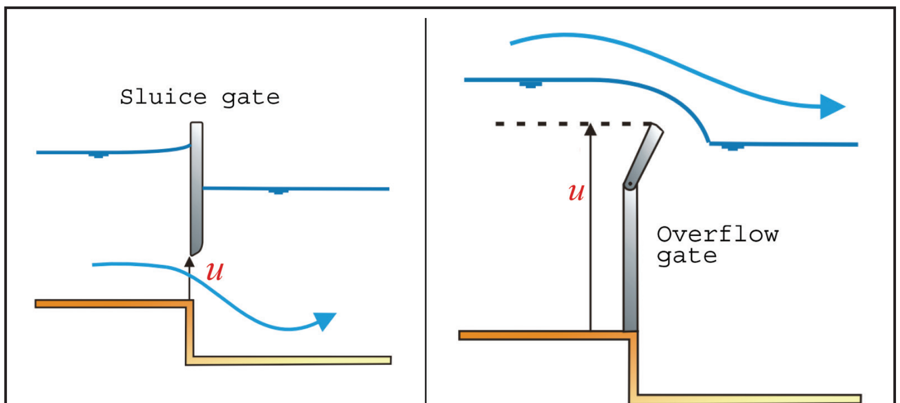


Figure 6. *A sluice gate and spillway, control devices that Coron and his colleagues use to rapidly stabilize the water level on the Sambre and Meuse rivers. Figure credit: Georges Bastin.*

1 https://youtu.be/gZgDWTtp2qs
2 https://youtu.be/iiEvqpqQ4fU

3 https://youtu.be/9y_qniyxaig

4 https://www.pathlms.com/siam/courses/4988/sections/7432

5 John Burns, a student of Reid's and the 2010 recipient of the Reid Prize, provided a personal account entitled "William T. and Idalia Reid: His Mathematics and Her Mathematical Family" as his Reid Lecture. A PDF of his lecture and slides with synchronized audio are available from SIAM at https://www.pathlms.com/siam/courses/3609/sections/5154.

Demystifying Chance: Understanding the Secrets of Probability

Ten Great Ideas about Chance. *By Persi Diaconis and Brian Skyrms. Princeton University Press, Princeton, NJ, November 2017. 272 pages, $27.95.*

BOOK REVIEW
By James Case

For the better part of a decade, Persi Diaconis and Brian Skyrms taught a course at Stanford University on the history, philosophy, and common foundations of probability and statistics. With the passage of time, they realized that the story they were telling would likely be of interest to a larger audience. Thus, *Ten Great Ideas about Chance* was born.

As the title suggests, the book consists of 10 chapters exploring 10 significant ideas about chance. An appendix offers a tutorial on probability and extensive chapter notes, an index, and an "annotated select bibliography." The latter comprises 10 numbered sections listing 41 seminal books and papers, with brief commentary on each. The entire book can be considered an extended digest of this list.

Some chapters—such as the fifth, concerning the mathematics of probability—are more or less obligatory in a book of this nature. After a few words about finite probability and a brief exposition of Borel and Cantelli's proof of the strong law of large numbers, the authors describe the sixth of Hilbert's 23 challenge problems. In this problem, Hilbert proposed that those physical sciences wherein mathematics—especially the theories of probability and mechanics—plays a significant role be placed on a sound axiomatic basis. He was apparently thinking of Ludwig Boltzmann's theory of gases, in which a swarm of hard spheres moves about in a rigid container; the spheres rebound off one another and the surrounding walls without losing momentum. Can one demonstrate, given a plausible prior distribution on the spheres' initial positions and momenta, that low-entropy states are likely to evolve into high-entropy states?

Little came of Hilbert's suggestion until 1933, when Andrei Kolmogorov published his groundbreaking book [1] on the foundations of probability theory. Kolmogorov did three important things: used measure theory to place probability on a firm mathematical foundation, formalized the previously nebulous concept of conditional probability, and proved an extension theorem that shows how an infinite-dimensional stochastic process can be built up from a consistent family of finite-dimensional probability spaces. His work led to an almost immediate flowering of probability theory that continues to this day.

Ten Great Ideas about Chance. By Persi Diaconis and Brian Skyrms. Courtesy of Princeton University Press.

Equally indispensable to Diaconis and Skyrms' purpose is a chapter on inverse inference, beginning with the question that concerned Reverend Thomas Bayes: after a coin of unknown bias has come up heads n times in $N > n$ trials, what are the odds that the probability p of its occurrence in a single subsequent trial lies within a given subinterval of [0,1]? Bayes solved this problem on the assumption that p is equally likely to lie anywhere in the unit interval before trials begin. Laplace later revisited Bayes' problem and arrived at his famous "rule of succession," $p_{est} = \frac{n+1}{N+2}$. For large n and N, this scarcely differs from the naïve estimate $\frac{n}{N}$. Modern critics have argued that, for an ordinary-looking coin, probabilities near the middle of [0,1] seem more likely than those at either extreme. Indeed, postulating a prior beta distribution $B(x;\alpha,\beta)$ on p shows that the same n heads in N trials leads to an updated beta distribution with parameters $\alpha+n$ and $\beta+N-n$. Hence, $p_{est} = \frac{n+\alpha}{N+\alpha+\beta}$, which again approximates $\frac{n}{N}$ for large n and N.

Bayes' theorem may present a valid rebuttal to philosopher David Hume's 1748 essay, "An Enquiry Concerning Human Understanding," which criticized conclusions drawn from records of past events. As investment advisors are honor-bound to warn potential customers that "past performance need not be indicative of future results," predictions predicated on the assumption that the future will resemble the past are inherently risky and should not be acted upon without prior assessment of this source of risk. Hume also pointed out that randomness does not exist in nature (or did not seem to until quantum phenomena came to light) because in his day people believed that knowledge of Newton's laws, together with the positions and momenta of every particle in the universe at one single instant, determined the entire future.

Another obligatory chapter concerns frequentism—the leading alternative to Bayesian inference—and the related notion that probability is a state of mind rather than a physical attribute observable only through repeated trials. The authors describe attempts by John Venn in 1866 and Richard von Mises in 1919 to base a coherent theory of probability on the premise that frequency testing alone can determine probabilities. Venn and Mises also tried to expose the fallacy in Johann Bernoulli's argument that his weak law of large numbers makes it possible to determine the chance that a specific outcome will be forthcoming on a single trial, given the results of a sufficient number of previous trials. The authors concede that it is a subtle fallacy, yet one that notables like Borel, Kolmogorov, Paul Levy, and Andrey Markov have failed at times to avoid.

See **Demystifying Chance** *on page 9*

Focusing on Nephroids

MATHEMATICAL CURIOSITIES
By Mark Levi

The bright strip in Figure 1 is illuminated by the incoming parallel beam reflected from the inner surface of the cup. The rays do not focus at a single point, as they would if the wall were parabolic. Instead, they "focus" at a curve, with the cusp at the focus of the osculating parabola.

Figure 2 sheds some light on the situation: the density of reflected rays spikes at the envelope, referred to as the *caustic*. This explains the caustic's brightness.

Remarkably, this caustic is an epicycloid — more precisely, the path of a particle on the rim of the wheel rolling without sliding on another wheel, with a 1:2 ratio of radii (see Figure 3). Because of its vaguely kidney-like shape, this epicycloid is referred to as a *nephroid* (kidney=$\nu\varepsilon\varphi\rho o$).

The precise statement is the following. *For every ray from a pencil of parallel rays striking the inside of a circular mirror of radius R, the reflected ray is tangent to the nephroid generated by rolling a circle of radius $R/4$ on the stationary circle of radius $R/2$ concentric with the mirror.*

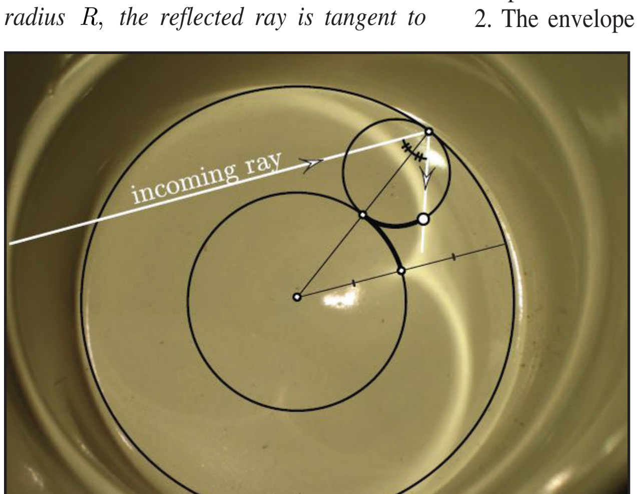


Figure 1. *Imperfect focusing creates a caustic.*

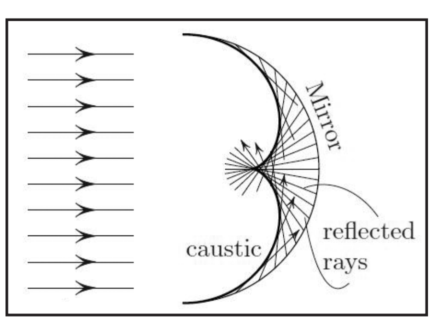


Figure 2. *The density of reflected rays spikes at the caustic. This caustic is half of the nephroid in Figure 3.*

The cusps of this nephroid lie on the ray passing through the center of the mirror.

To restate this result, imagine walking around the circle with constant angular velocity 1 and twirling a baton in the horizontal plane with angular velocity 2. The envelope of the resulting family of lines is a nephroid.

Proof of the Claim

Consider the nephroid generated by rolling the circle, as seen in Figures 3 and 4; T is the point tracing out the nephroid. We must prove that PT is the reflected ray, i.e., that

$$\angle APO = \angle OPT, \quad (1)$$

and that PT is tangent to the nephroid.

Note that $\text{arclength}(CS) = \text{arclength}(CT)$ due to the non-slip condition, so that

$$\angle COS = \frac{1}{2}\angle CQT,$$

because the radii are in the 1:2 ratio. In turn, $\angle CQT = 2\angle CPT$, since these angles subtend the same arc. This proves (1).

Why is TP tangent to the nephroid? As the smaller circle rolls on the larger one, the velocity $v_T \perp CT$, because C is the rolling wheel's instantaneous center of rotation. But $TP \perp CT$, since CP is a diameter. Thus, $v_t \parallel TP$; TP is indeed tangent. □

On a related note, if the source of light lies on the circle, then the resulting caustic is a cardioid, i.e., the epicycloid generated by the circle rolling on another circle of equal radius (see Figure 5). This proof is the same as the one above.

To conclude, here is a small challenge: show that the length of the thick line in Figure 5 is independent of the choice of B, namely

$$AB + BT + TC = \frac{8}{3}R, \quad (2)$$

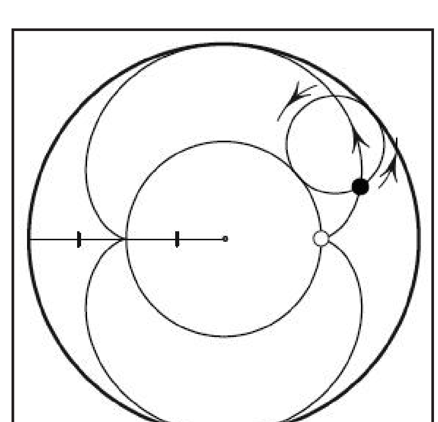


Figure 3. *The nephroid: an epicycloid with a 1:2 ratio of radii.*

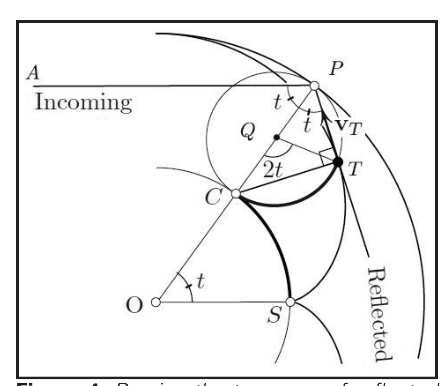


Figure 4. *Proving the tangency of reflected rays to the nephroid, which is traced by point T on the rolling wheel. The starting position of T is S.*

R being the radius of the mirror. Assuming that (2) holds, setting $B = A$ causes the first two terms to vanish, and we conclude that the cardioid's length is $16R/3$.

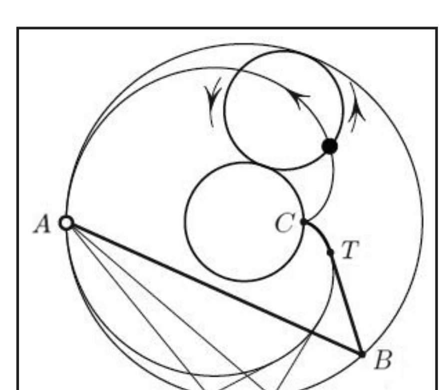


Figure 5. *The cardioid—an epicycloid with 1:1 ratio—is the caustic created by the source of light lying on the circular mirror.*

The figures in this article were provided by the author.

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

Quantifying, Reducing, and Repurposing Wasted Food

M3 Challenge Problem Writers Discuss this Year's Question

The MathWorks Math Modeling (M3) Challenge—a competition that encourages participants to use critical thinking and computational methods to quantify open-ended, real-world problems—poses prominent, multi-faceted issues to U.S. high school juniors and seniors, who compete on teams of three to five. To make the contest accessible to all participants regardless of their prior exposure to mathematical modeling, SIAM is constantly on the lookout for interesting problems that enable teams to find data and implement approaches with varying degrees of simplicity or complexity. Exposing thousands of future leaders and innovators to important topics in applied mathematics is among SIAM's priorities.

Judges for the contest—who numbered about 130 this year—anticipate the commendable creativity and insights displayed by students in their solution papers, motivated by a passion for mathematical applications and a shot at winning part of the $100,000 in scholarship awards funded by MathWorks. 912 teams—comprised of 4,175 high school juniors and seniors—submitted solution papers in the 2018 Challenge, each vying for one of 40 monetary prizes.

The 2018 Challenge problem focused on food insecurity. It posited that "ugly," uneaten, leftover, and excess food waste from households, cafeterias, restaurants, and grocery stores is not "trash," but rather wasted food. Could a community feed those experiencing food insecurity by thoughtfully repurposing food squandered by others? How do we quantify the amount of food wasted annually? How can we redirect wasted food to benefit those in need? The problem posed a few specific household scenarios—single parent, family of four, etc.—on which to demonstrate models (though students could certainly utilize their own household structures), and asked for model-based strategies to repurpose the most food for the least cost.

Karen Bliss (Virginia Military Institute), Katie Kavanagh (Clarkson University) and Ben Galluzzo (Shippensburg University), who comprise the Challenge's problem development committee, crafted the 2018 problem, entitled "Better ATE than Never: Reducing Wasted Food" (see sidebar). Since 2012, the committee has formulated or refined submitted problem ideas, enabling the contest to demonstrate the use and value of mathematics in computation-based decisions while bringing visibility and novel thought to relevant, timely, and newsworthy topics. To ensure that the questions are authentic and realistic, the group often recruits and collaborates with professionals with expertise in the underlying area. The U.S. Department of Agriculture's Economic Research Service weighed in on this year's problem.

We asked problem development committee members: what motivated this year's problem on food waste?

Karen Bliss

We felt very strongly that this year's question should have a local tie-in, so participants felt connected to the topic. In addition to understanding the mathematics of it all, we wanted students to think about how they could apply modeling to a community to which they belong right now, be it their school, town, or state. We are hoping that students see how they can utilize applied mathematics in very concrete ways to assist those in need around them.

I was surprised by the volume of data available regarding food production, food consumption, and the severity of food insecurity. It became clear that researchers are actively studying this very worthy topic, and it will certainly be exciting to see how teams tackle the complex questions we posed. We wanted to introduce available data with the problem, since last year's Challenge question—which included data for students to grapple with as they did their modeling—was well received.

While modeling is certainly distinctly different from statistics and/or simply data-fitting with appropriate lines and curves, we hope that the provided data gave students a sense of what factors they might consider when building their models.

I cannot wait to hear how many teams show their work to their schools, towns, etc., to raise awareness for this cause and inspire action. That will be incredibly powerful.

Katie Kavanagh

My mathematical area of research is optimization, so my mind is programmed to look for ways to increase efficiency. I also have a decent-sized garden, so I care deeply about the time, labor, and resources that go into planting a seed and fostering it to grow into nourishment. Furthermore, I live in a rural, poverty-stricken region of upstate New York, where over 80 percent of the children in most of our school district are eligible for reduced or free lunch; it is the only daily meal some of them get. All of these factors motivated me to focus a problem on this critical subject.

Mathematically, modeling human behavior is complex and interdisciplinary, with a wide range of challenges. What factors and human characteristics lead one person to toss something out and another to make soup from their leftovers? What is the food industry's tipping point to produce the quantity that is actually needed and distribute it accordingly, so that the least possible amount winds up in a landfill? Luckily, certain initiatives provide data that could be useful in answering these questions. Mathematics can play a significant role in quantifying the amount of unconsumed food versus the amount that is actually grown, and designing and analyzing strategies to repurpose these resources. These are the driving ideas behind these questions; how much is wasted, how many are hungry, and can we close the gap? How do we model human behavior when it comes to food choices? How can we make changes?

Although it seems like this is a problem for agricultural economists and social scientists, I genuinely believe that talented and motivated high school students will come up with some innovative approaches. They see the behavior of their peers in the cafeteria and observe how their family prepares meals and shops for groceries. They may even stop and think about these actions the next time they decide to toss their brussels sprouts in the garbage. Students have access to a wide range of data and computational tools with which to create meaningful mathematical models, but their personal experiences and communities will also play a key role in their results. It may not even be upper-level mathematics content that leads to the most sophisticated model, though a successful solution will require innovation and a willingness to dive into the problem.

I am so excited to see what they come up with, and hope that they feel as passionately about the topic as I do.

Ben Galluzzo

The dining hall at Shippensburg University went "tray-less" a number of years ago, and noted a significant decrease in uneaten food as a result. Likely due to my knowledge of this data, as well as historical reminders that one should finish a meal with a "clean" plate, I assumed that "leftover" or uneaten prepared food accounted for most food waste.

However, two events brought the magnitude of food waste into perspective, so much so that they convinced me this was a problem worth investigating. First, Katie shared a National Geographic article[1] that highlights the overwhelming amount and types of perfectly fresh, unprepared food that is squandered via the supply chain. The reported numbers—2.9 trillion pounds of food, or about a third of global food production—are staggering. Soon after reading this article, I spent a day volunteering at the Central Pennsylvania Food Bank, sorting produce for redistribution to agencies across the region that provide assistance to food-insecure individuals. It was simultaneously surprising and disturbing to see and touch hundreds of pounds of completely fresh, remarkably "normal" produce—such as green peppers, potatoes, and eggplants—that had been identified as unsellable to the general public and, without intervention, would have been thrown away. Following this revelation, a conversation with Food Bank employees about their efforts to redistribute food, including produce and nonperishable items found in their warehouse, reinforced to me the nontrivial, resource-heavy work that they perform on a daily basis. I left this experience motivated to find more data on the topic of food waste and truly interested in exploring ways to redistribute food.

Who better to tackle this issue than motivated, talented high school students who will likely face firsthand some of the consequences of food insecurity in the decades to come? I look forward to reading about how participants choose to use mathematical modeling to approach this important topic.

[1] https://www.nationalgeographic.com/magazine/2016/03/global-food-waste-statistics/

Better ATE than Never: Reducing Wasted Food

The Food and Agriculture Organization of the United Nations reports that approximately one third of all food produced in the world for human consumption every year goes uneaten. As an example, perfectly good produce that is considered misshapen or otherwise unattractive is regularly discarded before reaching your grocery store shelves. The problem is even more pronounced in the U.S., where the Environmental Protection Agency (EPA) estimates that more food reaches landfills and incinerators than any other single material in our trash. Uneaten food also wastes resources (water, fertilizer, pesticides, land, etc.) used in food production. At the same time, it has been estimated that over 42 million Americans are food-insecure and could take advantage of all of this squandered food, frequently described as "wasted food."

1. Just Eat It! Create a mathematical model that a state could use to determine if it could feed its food-insecure population using the wasted food generated in that state. Demonstrate how your model works for Texas; you may choose to use provided data.

2. Food Foolish? Personal choices when it comes to food consumption primarily occur at the grocery store, school cafeteria, restaurants, and at home. Create a mathematical model that can be used to determine the amount of food waste a household generates in a year based on their traits and habits. Demonstrate how your model works by evaluating it for the following households (provided data may be helpful):

- Single parent with a toddler, annual income of $20,500
- Family of four (two parents, two teenage children), annual income of $135,000
- Elderly couple, living on retirement, annual income of $55,000
- Single 23-year-old, annual income of $45,000.

3. Hunger Game Plan? Communities are starting to recognize and address the opportunities associated with repurposing potentially wasted food. Think of a community that you belong to (your school, town, county, etc.) and use mathematical modeling to provide insight on which strategies they should adopt to repurpose the maximal amount of food at the minimum cost. In particular, quantify the costs and benefits associated with your strategies.

Your submission should include a one-page executive summary with your findings, followed by your solution paper — for a maximum of 20 pages. If you choose to write code as part of your work, please include it as an appendix; those pages will not count towards your 20-page limit.

Interested in seeing more? View archives dating back to 2006 of each year's Challenge problem, outstanding solutions, judge perspectives, etc., on the Challenge website: **http://M3Challenge.siam.org**.

SIAM Review's Top-downloaded Paper Spotlights a Rapidly-growing Field

By Mark Newman

On the occasion of SIAM Review's 60th volume, the author of the journal's most popular article offers insight as to why the paper continues to spark so much interest.

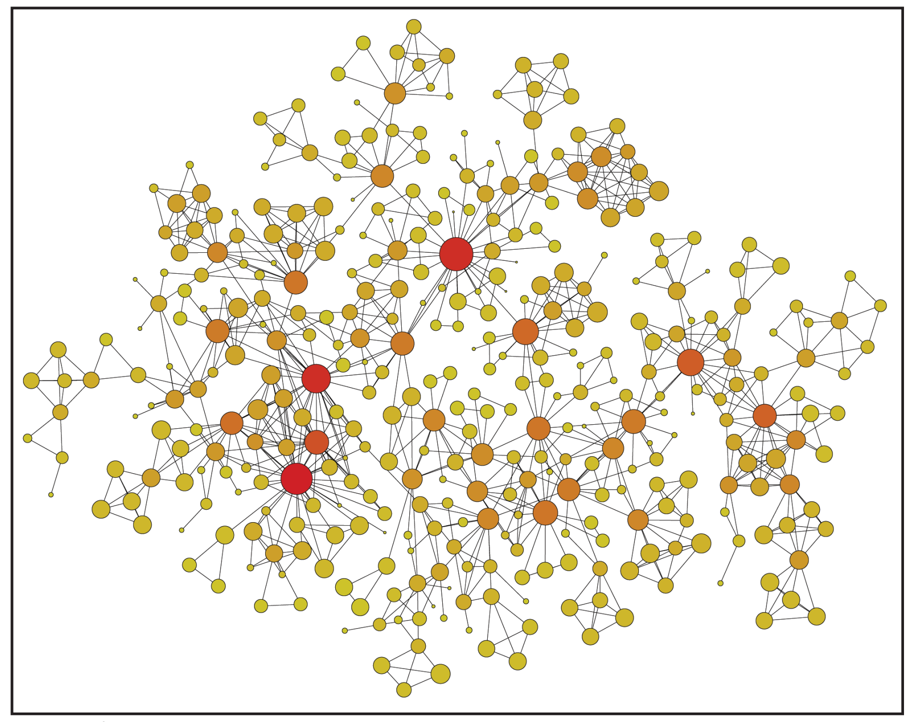
Many systems of scientific significance can be represented as graphs or networks — sets of nodes or vertices joined in pairs by edges. Examples include the internet; the World Wide Web; social, professional, and personal networks; road, rail, and airline routes; metabolic networks; food webs; and the power grid. In each of these cases, the network's structure can substantially impact system behavior. For instance, the flow of data traffic online depends strongly on the internet's network topology. How long will it take for data to travel from one part of the network to another? Are there bottlenecks or weak points in the structure? Would certain changes improve the system's performance or stability? The answers to these questions depend on the specific shape that the network takes.

Mathematicians have long studied graphs and networks in the context of graph theory, a branch of discrete mathematics that has yielded many beautiful formal results about network structure. Recent research, however, differs from traditional graph theory in its focus on the structure and properties of empirical networks as they appear in the real world. As a new area of applied research—sometimes dubbed "complex networks" or "network science"—this field has been driven in part by the increasing ubiquity of detailed data sets describing network structures across a range of different areas of science and technology, as well as the widespread availability of inexpensive and powerful computers with which to analyze them. It focuses on the development of novel mathematical theory and methods to analyze, quantify, and understand real-world networks.

Beginning in the 1990s, research on complex networks quickly identified a number of central issues important to the understanding of network phenomena that are still relevant today. These issues include the following: the construction and solution of formal models of network structure, such as random graphs and models of network growth; metrics that quantify specific structural features of networks, like path lengths, correlations, clustering coefficients, subgraph densities, and centrality measures; methods for quantifying large-scale structure in networks, particularly community structure; spectral graph theory and random matrix methods; networks' resilience to failure or attack; and processes taking place on networks, such as the spread of diseases in human populations or the flow of information across the internet.

By 2003, the field's focus had coalesced to the point where a survey of the mathematical developments was needed, and *SIAM Review* invited me to contribute an article.[1] The timing was ideal, coinciding with rapidly-increasing interest in networks across the mathematical sciences, and the article received a record number of citations in the years following its publication.

The field of network science has since grown to encompass thousands of researchers, with new papers appearing every day. Even after 15 years, the topics covered in the original review garner a significant amount of research attention, and the paper continues to be highly cited. But many new developments have emerged as well, including the study of dynamic networks (those that change over time), the development of new algorithmic and analytic methods for network data (including statistical inference and spectral methods), the study of multilayer and multiplex networks (those with multiple different types of edges), and theories of dynamical systems and processes occurring on networks (such as flow processes, synchronization, and cascading dynamics).



A scientific collaboration network. The nodes represent scientists and the links represent collaborations among them. Image credit: Mark Newman.

The field remains extremely active, with a number of new journals devoted to network topics and numerous conferences attracting large numbers of researchers, including the SIAM Workshop on Network Science, held each year in conjunction with the SIAM Annual Meeting. This year's workshop[2] will take place on July 12 and 13 in Portland, Ore.

Table 1 displays the 10 most-downloaded *SIAM Review* articles, all of which are freely accessible through the end of the year. This information is also available online.[3]

Mark Newman received a Ph.D. in theoretical physics from the University of Oxford and has held positions at Cornell University, the Santa Fe Institute, and the University of Michigan. He is currently the Anatol Rapoport Distinguished University Professor of Physics at Michigan and a professor in the university's Center for the Study of Complex Systems. His research focuses on mathematical theory and methodology surrounding the structure and function of empirically-observed networks, such as computer networks, social networks, and biological networks.

[1] http://epubs.siam.org/doi/abs/10.1137/S003614450342480

[2] http://www.siam.org/meetings/ns18/

[3] http://epubs.siam.org/page/sirev_celebrates_60_volumes

Rank	Article	Author(s)	Year	*SIAM Review* Section
1	The Structure and Function of Complex Networks	Mark E. J. Newman	2003	Survey and Review
2	Tensor Decompositions and Applications	Tamara G. Kolda and Brett W. Bader	2009	Survey and Review
3	An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations	Desmond J. Higham	2001	Education
4	Modeling and Simulating Chemical Reactions	Desmond J. Higham	2008	Education
5	Time-Frequency Analysis of Musical Instruments	Jeremy F. Alm and James S. Walker	2002	Education
6	Power-Law Distributions in Empirical Data	Aaron Clauset, Cosma Rohilla Shalizi, and Mark E. J. Newman	2009	Survey and Review
7	The Discrete Cosine Transform	Gilbert Strang	1999	Education
8	Barycentric Lagrange Interpolation	Jean-Paul Berrut and Lloyd N. Trefethen	2004	Education
9	The Mathematics of Atmospheric Dispersion Modeling	John M. Stockie	2011	Education
10	Theory and Applications of Robust Optimization	Dimitris Bertsimas, David B. Brown, and Constantine Caramanis	2011	Survey and Review

Table 1. *The 10 most downloaded SIAM Review articles.*

Demystifying Chance

Continued from page 7

The first of the book's 10 great ideas is the simple realization that chance can be measured. The authors reference Gerolamo Cardano's advice to gamblers and Jacob Bernoulli's correspondence with Blaise Pascal, Pierre de Fermat, and Christiaan Huygens, along with his proof of the weak law of large numbers, as early evidence of this fact. To be fair, the ancient Greeks and Romans were well aware of chance. Yet they tended to attribute the outcomes of wars, races, courtships, and other contested events to interventions of the gods or deeds of the so-called "Fates," depicted as a trio of women. It remains a mystery that thinkers as perceptive as Euclid, Plato, and Archimedes never enunciated a law of large numbers or formulated a theory of discrete probability.

Diaconis and Skyrms' second great idea is, unsurprisingly, that one can infer probabilities in situations where *a priori* estimates are either unavailable or unreliable. For example, if a coin turns up 17 heads in 50 tosses, it is only natural to suppose that (i) the coin is unfair and (ii) the likelihood of a head on the next toss is closer to $1/3$ than $1/2$. The authors cite Bruno de Finetti, Leonard "Jimmie" Savage, and Frank P. Ramsey as developers of the intuitive notion of subjective probability. The basic idea, which differs from the older concept of frequentism, is that probabilities can be inferred from a "coherent set of beliefs" concerning the possible outcomes of a particular chance event, such as a horse race or boxing match.

A set of beliefs is coherent if it is impossible to construct a "Dutch book" predicated on them. A Dutch book is a collection of wagers with positive overall expectation. For instance, if both entries in a two-horse race go off as 2-1 favorites, a Dutch book could consist of a $1 bet on each horse. The winning ticket would then return $3 while the loser would return nothing. The resulting $1 gain is an expectation rather than a guarantee, since a dead heat would return only the $2 held by the track in escrow.

Amos Tversky and Daniel Kahneman advanced an equally great idea: that people are remarkably inept in their responses to chance events. Chapter three features a careful analysis of the famous Allais paradox in light of Savage's axioms of rational decision-making. Economist Maurice Allais asked a number of respondents to choose between payoff schemes A and B in two quite similar lotteries. Whereas respondents constrained by Savage's seemingly self-evident "axiom of independence" would choose B in both cases, Allais found that many flesh-and-blood respondents—including both Diaconis and Skyrms—chose B in the second case but A in the first. The experiment has been repeated multiple times worldwide, with quite similar results.

Shortly after Allais published his findings, Daniel Ellsberg—who later released the Pentagon Papers and became an activist against the Vietnam War—proposed a series of problems intended to illustrate the difference between choices involving *risk*, where objective probabilities are known, and *uncertainty*, where they are not. Since then, Kahneman and Tversky have described numerous situations in which the psychology of chance conflicts with its logic. The authors suggest that training in decision theory might improve real-world outcomes, especially in medical decision-making.

The book's final four chapters are at once more mathematically-challenging and philosophically-probing than their predecessors. The final chapter on inference seems to summarize all that modern scholarship has added to Hume's timeless classic on human understanding.

Ten Great Ideas about Chance is not a book to be read in bed at night. It should be attacked with paper and pencil at hand, and a determination to backtrack early and often. The extra effort will prove rewarding to almost any reader.

References

[1] Kolmogorov, A.N. (1950). *Foundations of the Theory of Probability*. New York, NY: Chelsea. (Original work published 1933).

James Case writes from Baltimore, Maryland.

NSF Releases Details of Proposed Fiscal Year 2019 Budget

By Eliana Perlmutter and Miriam Quintal

On February 28, the National Science Foundation (NSF) published its detailed budget request for fiscal year (FY) 2019. Outlining agency priorities and proposed funding levels as directed by the Trump administration, the request came weeks after the release of the president's government-wide FY 2019 federal budget proposal. The White House had planned to propose deep cuts—almost 30 percent below FY 2017 levels—to the NSF. However, due to increased spending caps in the new budget agreement for FY 2018 and FY 2019, the administration added $2.2 billion back into the NSF budget request, allowing essentially flat funding overall and a 2.4 percent increase to $6.15 billion for the Research and Related Activities account that funds all NSF research directorates. The president's budget request release discloses agency and administration priorities, and kicks off the appropriations process to fund the government for FY 2019. Congress is ultimately responsible for determining funding levels for individual agencies. FY 2019 will start on October 1, 2018, although final appropriations levels are not expected until after the completion of midterm elections this fall.

The NSF continues to focus on its 10 Big Ideas for Future NSF Investments,[1] launched by director France Córdova in May 2016. The NSF will provide $342 million for the ideas, which address exciting scientific challenge areas, thus enabling growth and new programs across all ten domains. Six of the ideas—including Harnessing the Data Revolution (HDR), the Future of Work at the Human-Technology Frontier (FW-HTF), Navigating the New Arctic, Windows on the Universe: the Era of Multi-Messenger Astrophysics (WoU), the Quantum Leap: Leading the Next Quantum Revolution (QL), and Understanding the Rules of Life: Predicting Phenotype—will receive $30 million under the budget request. While individual directorates will hold funding for each Big Idea, cross-directorate working groups will continue to lead the initiatives and determine specific thrusts and investments.

$60 million in funding is included for two new convergence accelerators related to HDR and FW-HTF. These accelerators will seek partnerships with other agencies, industry, foundations, and international organizations to support translational research, testbeds, infrastructure access, and workforce considerations. The other four Big Ideas supported under the budget request center on NSF process improvements to enable science and engineering advancement. For example, mid-scale research infrastructure will receive $60 million under the Office of Integrative Activities to facilitate mid-scale funding for projects across science and engineering disciplines. Additionally, the NSF proposes $16 million for growing convergence research to identify compelling convergent research challenges and fund exploratory science, engineering, and workforce efforts to tackle these challenges.

Many programs will see reductions under this budget proposal, including several Obama-era initiatives already slated to wind down, the Graduate Research Fellowship Program, Faculty Early Career Development (CAREER) awards, and multiple education programs. Apart from major cross-cutting initiatives and details about certain STEM programs, the budget request provides no information about whether or how individual divisions will apply reductions across their core activities. The Directorate for Mathematical and Physical Sciences (MPS) will face a 1.3 percent reduction from its FY 2017 level, for a total funding level of $1.345 billion. The Division of Mathematical Sciences (DMS) will be down 6.3 percent to $219 million. All of the disciplinary divisions in the MPS are expected to decrease from their FY 2017 levels by between five and 8.2 percent, while the Office of Multidisciplinary Activities (OMA) is projected to grow 197 percent to $103 million. This growth reflects the assumption that the OMA will hold the funding for QL and WoU, two of the Big Ideas. Elsewhere in the foundation, the Directorate for Computer and Information Science and Engineering will be down 1.1 percent from FY 2017 to $925 million, and the Office of Advanced Cyberinfrastructure will drop 5.9 percent to $210 million. Despite the proposed decreases, NSF leadership has indicated their continued commitment to a robust core research portfolio.

While the president's budget request recommends flat funding for the NSF overall in FY 2019, the research advocacy community is pushing for robust growth to $8.45 billion. SIAM will continue to participate in these efforts and ensure that Congress is aware of both the NSF's importance to the applied mathematics and computational science community, and the need to guarantee NSF funding at levels permitting strong investment in the Big Ideas while also protecting core activities.

View the complete NSF FY 2019 budget request online.[2]

Eliana Perlmutter is a Legislative Research Assistant and Miriam Quintal is SIAM's Washington liaison at Lewis-Burke Associates LLC.

[1] https://www.nsf.gov/news/special_reports/big_ideas/

[2] https://www.nsf.gov/about/budget/fy2019/toc.jsp

The High-Performance Geometric Multigrid: An HPC Performance Benchmark

By Samuel Williams, Mark F. Adams, and Jed Brown

The High-Performance LINPACK (HPL) Benchmark is the most widely-recognized metric for ranking high-performance computing (HPC) systems. It rose to prominence in the early 1990s, when its predicted ranking of a system correlated with the system's efficacy for full-scale applications. Computer system vendors sought designs that would increase HPL performance, which in turn improved overall application behavior.

Unfortunately, this is no longer the case; in fact, the opposite is now true. Although the HPL Benchmark continues to be an effective proxy for applications based on dense linear algebra, it has lost its proficiency as a proxy for many applications relevant to the HPC community. HPL rankings of computer systems, which utilize work-optimal algorithms with high bandwidth and low latency requirements, are not well-correlated to real application performance nowadays. Motived by HPL's increasing inapplicability, the High-Performance Geometric Multigrid (HPGMG) incorporates machine sensitivities that correlate well with the sensitivities of HPC applications.

HPGMG complements both HPL and the new High-Performance Conjugate Gradients (HPCG) Benchmark[1] [2], with more stress on the memory system and network fabric than HPL and HPCG respectively. The TOP500 list is currently adding new rankings for HPCG and HPGMG to complement the venerable HPL.

HPGMG Design Principles

The following design principles of HPGMG are discussed in the HPGMG 1.0 whitepaper [1]:

> A benchmark must reflect improvements to computer systems that benefit our applications, and is essential for documenting future improvements to HPC systems. The metric must be designed so that, as we optimize metric results for a particular platform, the changes will also lead to performance improvements realized by real applications. Any new metric we introduce must satisfy a number of requirements:
>
> • Accurately reflect the characteristics of contemporary high-efficiency algorithms.
>
> • Accurately reflect the principle challenges of future hardware design — a balanced combination of memory bandwidth, interconnect performance (both for small and large messages), computational performance, and reliability. It should not be possible to "cheat" the benchmark by over-provisioning the hardware in any one of these areas. A machine designed for the sole purpose of performance on our metric should result in a "pretty good" machine for scientific and engineering applications.
>
> • The absolute improvements in this benchmark should ultimately be reflective of performance improvements realizable in real applications, which are occurring at a much slower rate than improvements in peak FLOPs.
>
> • It must be able to scale through many orders of magnitude improvement of hardware storage

See **Geometric Multigrid** *on page 12*

[1] https://sinews.siam.org/Details-Page/the-high-performance-conjugate-gradients-benchmark

Rank	Site	System	10^9 DOF/s h, 2h, 4h	Parallelization MPI	OMP	ACC	DOF per Process	TOP500 Rank
1	National Supercomputing Center in Wuxi *(China)*	**Sunway TaihuLight** – Sunway MPP, SW26010 260C 1.45GHz, Sunway, NRCPC	**1036** 565 163	131072	1	1	32M	1
2	Department of Energy/Office of Science/Lawrence Berkeley National Laboratory/National Energy Research Scientific Computing Center *(USA)*	**Cori** – Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect, Cray Inc.	**859** 376 87	65536	8	0	16M	8
3	Department of Energy/Office of Science/Argonne National Laboratory *(USA)*	**Mira** – BlueGene/Q, Power BQC 16C 1.60GHz, Custom interconnect, IBM	**500** 313 107	49152	64	0	36M	11
		(baseline)	395 286 107	49152	64	0	36M	
4	Höchstleistungsrechenzentrum Stuttgart *(Germany)*	**Hazel Hen** – Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect, Cray Inc.	**495** 411 221	15408	12	0	192M	19
5	Department of Energy/Office of Science/Oak Ridge National Laboratory *(USA)*	**Titan** – Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x, Cray Inc.	**440** 163 38.9	16384	4	1	32M	5
		(CPU-only)	161 82.5 23.7	36864	8	0	48M	
6	King Abdullah University of Science and Technology *(Saudi Arabia)*	**Shaheen II** – Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect, Cray Inc.	**326** 287 175	12288	16	0	144M	20
7	Department of Energy/Office of Science/Lawrence Berkeley National Laboratory/National Energy Research Scientific Computing Center *(USA)*	**Edison** – Cray XC30, Intel Xeon E5-2695v2 12C 2.4GHz, Aries interconnect, Cray Inc.	**296** 246 127	10648	12	0	128M	78
8	Swiss National Supercomputing Centre *(Switzerland)*	**Piz Daint** – Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect, NVIDIA K20x, Cray Inc.	**153** 68.8 18.5	4096	8	1	32M	–
		(CPU-only)	85.1 62.6 24.7	4096	8	0	16M	
9	Cyberscience Center, Tohoku University *(Japan)*	**SX-ACE** – 4C 1GHz, IXS NEC	**73.8** 45.2 15.6	4096	1	0	128M	–

Table 1. *A compressed sampling of the High-Performance Geometric Multigrid (HPGMG) list of the world's largest supercomputers, as of November 2017.*

Understanding and Appreciating Mathematics and Statistics

April is Mathematics and Statistics Awareness Month

Both mathematics and statistics play a significant role in addressing many real-world problems, including climate change, disease, sustainability, the data deluge, and internet security. Research in these and other areas is ongoing, yielding new results and applications every day in fields such as medicine, manufacturing, energy, biotechnology, and business. Mathematics and statistics are important drivers of innovation in our technological world, where new systems and methodologies continue to become more complex.

We encourage you to participate in Mathematics and Statistics Awareness Month![1] Organize and host activities all throughout April; past events have included workshops, competitions, festivals, lectures, symposia, department open houses, math art exhibits, and math poetry readings. Share your activities on social media with the hashtag #MathStatMonth. Follow "MathAware" on Facebook and Twitter for more information.

Schools, universities, organizations, associations, and related interest groups are hosting a myriad of activities throughout April in honor of Mathematics and Statistics Awareness Month.

Mathematics and statistics are powerful tools. Thus, Mathematics and Statistics Awareness Month is important because these subjects enable our understanding of the world around us. In the realm of physics, mathematics allows scientists to study planetary orbits and send space probes on successful journeys that are hundreds of millions of miles in length. Using statistics, physicists can sift through quadrillions of collisions in the Large Hadron Collider and discover the mysterious Higgs boson. In engineering, these tools facilitate the design of steel buildings, bridges, and airplanes, and ensure their performance and robustness even before they are constructed. Mathematics and statistics help biologists decode the mysteries of human DNA and prevent disease. In business, researchers use these subjects to craft financial predictions, study economics, and forecast strategies to benefit entire nations. They are at the heart of computers, technology, electrical systems, and the engines that drive our cars and fly our planes. Perhaps Galileo said it best: "Mathematics is the language with which God has written the universe."

Mathematics is the science of patterns and statistics is the science of data. Learning these subjects is like learning to play an intricate, challenging, and exciting game. Every branch of mathematics has its own rules and we explore the possibilities within these rules, discovering new ideas and opportunities as we go. The rules of mathematics and statistics are the rules of logic; they existed long before we ever did.

Finally, let us not forget that mathematics and statistics are beautiful, extraordinary forms of art. They exhibit as much detail as a sculpture by Michelangelo, carry the universal resonance of a Beethoven symphony, elicit the serene beauty of a painting by Leonardo da Vinci, and reveal insights as deep and enduring as those in a Shakespearean play. Their powerful truths will continue to transcend time and shape the ideas of future generations.

— *Kelly Cline, Carroll College*

SIAM pulls together individuals who are not easy to convene — applied mathematicians and computational scientists come from many academic departments and hold numerous industrial job titles, few of which include the word "mathematician." The organization's diversity and interdisciplinary nature provide a space for people from these various backgrounds all over the world. As Mathematics and Statistics Awareness Month approaches, I appreciate SIAM's unique role among professional societies in bringing real-world applications of math to so many issues. These issues range from tumor growth to baseball, from modeling the dynamics of mosquito-borne diseases or drug-dosing schedules to predicting financial markets or the spread of Lyme disease, and ultimately involve the utilization of data, quantitative methods, and (frequently) technical computing software to bring better understanding and insight to big, important, open-ended topics that are essential to our world. SIAM encourages and supports research and education in the mathematical sciences via publications and conferences, which give professionals the opportunity to share their work with others in the community. It also aims to deliver the valuable work occurring in our fields to the public eye through its website, publicity and press releases, social media, YouTube videos, and even the presentation of relevant problems to thousands of high school students each year in the MathWorks Math Modeling Challenge. Bringing that visibility of computational sciences is really the crux of Mathematics and Statistics Awareness Month, so kudos for all of the good work being done and to SIAM for putting a very public lens on it.

— *Cleve Moler, MathWorks*

1 http://www.mathaware.org

New Division Director for the NSF's Division of Mathematical Sciences

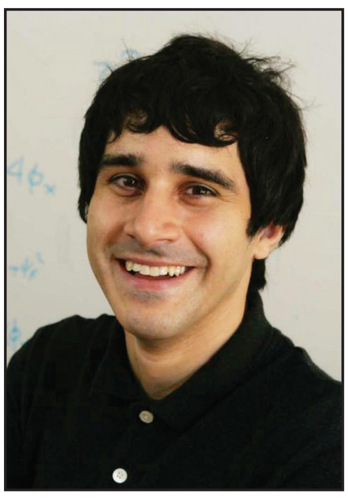
Juan Meza has been appointed the new division director of the National Science Foundation (NSF)'s Division of Mathematical Sciences (DMS). Meza served most recently as the dean of the University of California, Merced's School of Natural Sciences. His remarkable career spans over three decades, with leadership and management experiences in industry, academia, and national laboratories. He was a Distinguished Member of Technical Staff at Sandia National Laboratories and department head and senior scientist of high-performance computing research at Lawrence Berkeley National Laboratory. Meza was also a member of both the NSF's Advisory Committee for Mathematical and Physical Sciences and the Advisory Committee for Cyberinfrastructure, as well as multiple Committees of Visitors for the DMS. He has served on the board of directors for the National Academies of Sciences, Engineering, and Medicine's Board on Mathematical Sciences and its Applications, the Society for the Advancement of Chicanos/Hispanics and Native Americans in the Sciences, and the American Association for the Advancement of Science Council, representing the Section on Mathematics. Meza holds both a bachelor's and master's degree in electrical engineering, and a master's and Ph.D. in computational and applied mathematics, all from Rice University. He is a former member of SIAM's Board of Trustees.

Snow Business: Computational Elastoplasticity in the Movies and Beyond

By Joseph Teran

The following is a short introduction to the American Mathematical Society Invited Address, to be presented at the upcoming 2018 SIAM Annual Meeting (AN18) in Portland, Ore., from July 9-13. Look for feature articles by other AN18 invited speakers introducing the topics of their talks in future issues.

Over the past two decades, visual effects have come to rely on a wide range of numerical methods for partial differential equations (PDEs). Be it the crashing water waves in Disney's *Moana* or falling snow in *Frozen*, audiences now expect the computer-generated world to look and *move* like the real thing. The demand for realism is so high that it is impractical—or impossible—for animators to reproduce the dynamics of everyday materials like clothing, water, sand, snow, or hair without using the laws of physics that dictate their motion.



Joseph Teran, University of California, Los Angeles

The governing physics is expressed with PDEs derived from classical continuum mechanics (e.g., the Navier-Stokes equations for water). The PDEs are highly nonlinear and involve geometrically-complicated domains, like the upper body of the character in Figure 1a and the snow under Anna's feet in Figure 1b. Given these constraints, one can only solve the equations with numerical approximation and scientific computing. Techniques from computational fluid dynamics, like particle-in-cell [1-2, 4] and the finite element method for elastic solids [6], are now commonplace in the production of blockbuster movies for these reasons.

Many everyday materials behave elastically for a wide range of strains, but plastically upon approaching nonphysical stresses. Common examples include metals and granular materials like sand, snow, mud, and dirt. One can even describe frictional contact as a plastic constraint on states of stresses that arise during contact. In continuum mechanics, the Cauchy stress σ is defined by the relation between internal surfaces of contact with normals n and the contact force per unit area (or traction) $\mathrm{t}=\sigma\mathrm{n}=\mathrm{t}_\tau+(\mathrm{n}\cdot\sigma\mathrm{n})\mathrm{n}$. When the contact force must obey Coulomb friction, the tangential component t_τ of the force must be smaller in magnitude than a coefficient of friction c_F times the normal component $-\mathrm{n}\cdot\sigma\mathrm{n}$:

$$|\mathrm{t}_\tau|\leq -c_F\mathrm{n}\cdot\sigma\mathrm{n}. \tag{1}$$

For example, one can derive the Mohr-Coulomb and Drucker-Prager plastic yield conditions [5] for granular materials by simply applying this Coulomb friction condition to a hyperelastic constitutive model (see Figure 1b). We have recently shown that even clothing can be simulated from a continuum view, where Coulomb friction during contact places a constraint on the types of stresses that are physical (see Figure 1c). The material point method [8] is key to translating these continuum descriptions of plasticity physics into discretized approximations that one can use for visual effects. This technique is a generalization of the particle-in-cell approach [1, 2] to history-dependent materials, and is not used for a broad range of materials whose physics is naturally described by elastoplasticity.

I will discuss these aspects and more during my talk at the 2018 SIAM Annual Meeting.

References

[1] Brackbill, J., & Ruppel, H. (1986). FLIP: A method for adaptively zoned, particle-in-cell calculations of fluid flows in two dimensions. *J. Comp. Phys., 65*, 314-343.

[2] Harlow, F. (1964). The particle-in-cell method for numerical solution of problems in fluid dynamics. *Meth. Comp. Phys., 3*, 319-343.

[3] Jiang, C., Gast, T., & Teran, J. (2017). Anisotropic elastoplasticity for cloth, knit and hair frictional contact. *ACM Trans. Graph., 36*(4) 152:1-152:14.

[4] Jiang, C., Schroeder, C., Selle, A., Teran, J., & Stomakhin, A. (2015). The affine particle-in-cell method. *ACM Trans. Graph., 34*(4), 51:1-51:10.

[5] Klar, G., Gast, T., Pradhana, A., Fu, C., Schroeder, C., Jiang, C., & Teran, J. (2016). Drucker-prager elastoplasticity for sand animation. *ACM Trans. Graph., 35*(4),103:1-103:12.

[6] McAdams, A., Zhu, Y., Selle, A., Empey, M., Tamstorf, R., Teran, J., & Sifakis, E. (2011). Efficient elasticity for character skinning with contact and collisions. *ACM Trans. Graph., 30*(4), 37:1- 37:12.

[7] Stomakhin, A., Schroeder, C., Chai, L., Teran, J., & Selle, A. (2013). A material point method for snow simulation. *ACM Trans. Graph., 32*(4), 102:1-102:10.

[8] Sulsky, D., Chen, Z., & Schreyer, H. (1994). A particle method for history-dependent materials. *Comp. Meth. App. Mech. Eng., 118*(1), 179-196.

Joseph Teran is a professor in the Department of Mathematics at the University of California, Los Angeles.

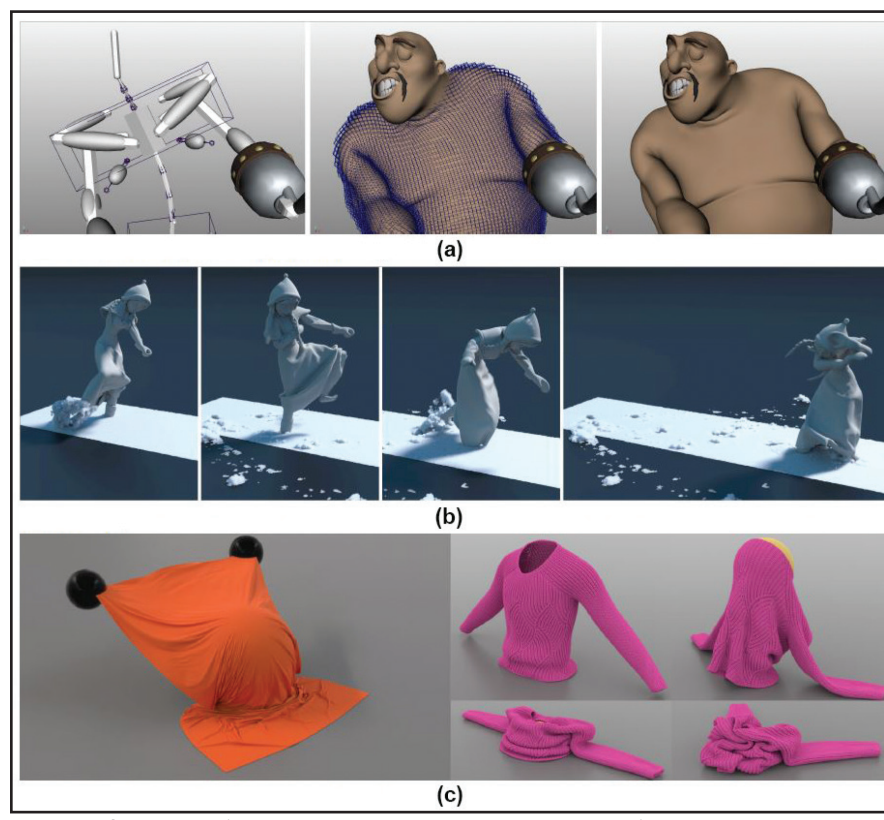


Figure 1. *Simulation of hyperelastic and plastic materials.* **1a.** *The finite element method simulates soft elastic tissue on a Disney movie character. The discrete simulation mesh is shown in blue. Image courtesy of [6].* **1b.** *The material point method simulates the snow under Anna's feet as an elastoplastic granular material in Disney's Frozen. Image courtesy of [7].* **1c.** *One can model frictional contact between layers of clothing with a discretized elastoplastic continuum model. Images courtesy of [3].*

Geometric Multigrid

Continued from page 10

capacity and performance — much as HPL has for the past three decades.

No one benchmark can provide an accurate proxy of any particular application, but we believe that one comprehensive benchmark has two advantages over the alternative: a weighted set or bag of (simple) benchmarks.

The weighted bag of benchmarks does allow for explicit definition of machine metrics, measurement of these metrics in the benchmark and applications, and fitting of internal parameters so as to provide the best proxy for an application of interest. This approach can be more accurate than a single benchmark for a particular application or specific workload. One comprehensive benchmark, like HPL or HPGMG, is easier to administer, define, and adjudicate for a ranking metric.

Though we seek a rational approach—with modeling and measurements—for benchmark design, models cannot perfectly measure machine effectiveness. Therefore, a benchmark that implicitly demands an effective machine by solving a fundamentally hard problem is desirable. HPGMG has been designed with these principles in mind.

HPGMG Design

HPGMG uses the non-iterative form of multigrid—full multigrid—which requires $O(N)$ flops, $O(N)$ bytes from memory, $O(N)$ bytes from cache, $O(N^{2/3})$ MPI bytes, $O(\log^2(N))$messages, and $O(\log^2(N))$ function calls/OMP overheads/CUDA kernels. For fourth-order accurate finite-volume of HPGMG (used for the ranking metric), this equates to about 1,200 flops and 1,200 bytes per DOF solved. As such, the arithmetic intensity (AI) is about 1.0 flop per byte. These measures are pretty strongly memory-bound, as most machines have AIs of five to 10 flops per byte, and thus incentivize higher memory bandwidth.

In practice (as observed with LIKWID), the L2 and L3 cache bandwidths are four times and two times higher than DRAM bandwidth. This stimulates a tapered cache hierarchy with progressively higher bandwidth when getting closer to the FPUs, i.e., one cannot just attach an FPU to an HBM stack, but must exploit locality for bandwidth filtering. The number of messages is polylogarithmic in the problem size. As such, architects are forced to drive down overheads when scaling up machines to avoid squandering performance. This fact differentiates HPGMG from both supercomputing and HPCG, which sends $O(1)$ messages (a truncated V-cycle). MPI injection bandwidth is concurrently linked to problem size and thus memory bandwidth, but grows more slowly.

The polylogarithmic nature of HPGMG also manifests in the overheads for function calls, OMP parallel regions, and device offloads. This incentivizes architects and software technologies to drive down overheads as they increase memory and compute capacity, lest they forgo the benefit of threading or acceleration.

We compile the HPGMG list of the world's largest supercomputers twice a year—with the metric equations solved per second—using a multigrid solve of a fourth-order accurate finite-volume discretization of the Laplacian. We published our first list at ISC High Performance 2016, and have continued releasing lists biannually. Table 1 (on page 10) depicts a selected and compressed sampling of the November 2017 list.

We encourage community participation and invite comments and contributions to SWWilliams@lbl.gov. Visit the HPGMG webpage[2] to learn more about the effort. Detailed analysis and current rank lists are also available online.[3]

References

[1] Adams, M.F., Brown, J., Shalf, J., Straalen, B.V., Strohmaier, E., & Williams, S. (2014). *HPGMG 1.0: A Benchmark for Ranking High Performance Computing Systems* (LBNL Technical report, LBNL-6630E). Lawrence Berkeley National Laboratory.

[2] Dongarra, J., Heroux, M.A., & Luszczek, P. (2016). High-performance conjugate-gradient benchmark: A new metric for ranking high-performance computing systems. *Int. J. High Perf. Comp. App., 30*(1), 3-10.

Samuel Williams received his Ph.D. in computer science from the University of California, Berkeley in 2008, and is a staff scientist in the Computational Research Division at Lawrence Berkeley National Laboratory (LBNL). His work focuses on hardware/software solutions that affect performance, scalability, productivity, energy efficiency, and analytical capabilities on emerging multicore and accelerator-based supercomputers. Mark F. Adams received his Ph.D. in civil engineering from the University of California, Berkeley in 1998, and is currently a staff scientist at LBNL. He worked at Sandia National Laboratories and Columbia University before joining the LBNL staff in 2013. Jed Brown received his Dr.Sc. from ETH Zürich in 2011. He worked at Argonne National Laboratory until joining the University of Colorado Boulder as an assistant professor in 2015.

[2] https://hpgmg.org/
[3] https://crd.lbl.gov/hpgmg